



Framework for Improving Access to Medical Libraries using
Mobile Technologies

Fatimah Alzahrani

The thesis submitted in the partial fulfillment of the requirement of the
University of South Wales/ Prifysgol De Cymru for the degree of
Doctor of Philosophy.

2019

Abstract

Health information are rapidly increasing every day due to the vast amount of medical researchers related to diseases, drugs and diagnosis. Currently, huge databases and medical websites provide medical related -information to the users. Despite the considerable number of these resources, physicians and medical information seekers still have problems finding relevant information due to the huge number of results returned by these methods. However, a number of attempts carried out to solved that problem by providing a specific medical information retrieval systems and medical applications those approach were well designed but also referred to number obstacles such as the large number of results disorganized of information, learning how to access catalogues, and learning resources Moreover, these methods were limited on some services and did not focused on the user needs. As result this study aims to introduce a generic framework that allows users to access medical information. The framework aims to provide personalized information to meet the user needs based on their profile the framework designed in order to address the problem of understanding the users' needs and customizing the results in order to get to most relevant results that meet the user needs .the framework designed based on the principle of existing the communication system IMS and adapt the layer architecture and present a four main layered module; Access layer, control layer, application layer, finally smart layer which is designed for the purpose of personalizing the results. The designed framework will validated via simulations on a discrete event simulator for estimating the performance measures such as the delay and throughput under heavy load.

Acknowledgments

First of all, thank God that assisted me during the research project for giving me the inspiration, patience, time, and strength to finish this work. This thesis was the result of the collective efforts of a number of important and valued people who directly or indirectly assisted and supported me during my doctoral studies. I owe my deepest gratitude and thanks.

I would like to begin by thanking professor Andrew Ware for his support, supervision, and patience also for giving me the freedom and flexibility to allow me to develop as an independent researcher, I would like also to thank professor Khalid Al Begain for his support since the beginning of this research, I would like to thank him also for pushing me to achieve the very best work possible for which I am eternally grateful. Thank are also due to the University of South Wales for supporting this work.

I owe my deepest gratitude and appreciation to my husband Bandar Alzahrani who was here at every step to support and encourage me and showed patience and understanding towards my involvement in my work. Many thanks and regards to my mother who never stopped praying and blessing for me to achieve the best and to finish my research work. I would like to express my greatest gratitude and thanks to my daughter and son who brought happiness into my life, my sisters and my brother without them, I would not have had the strength or means to complete my PhD without their encouragement, love and support.

Finally, I am very thankful to King Saud University in the Kingdom of Saudi Arabia for granting me this opportunity to study and achieve my PhD degree I also would like to express my sincere thanks for the Saudi Arabian Cultural Bureau in London for the continued and generous support that I had during my studies in the UK.

Table of Contents

Abstract	i
Acknowledgments.....	ii
Table of Contents.....	iii
List of Figures	ix
List of Tables	xii
List of Abbreviations.....	xiii
Chapter 1: Introduction	1
1.1 Introduction.....	1
1.2 Research Problem and Motivation	3
1.3 Research Motivation.....	5
1.4 Research Objectives	7
1.5 Scope of the Study.....	8
1.6 Research Methodology	11
1.6.1 Designing the Framework	11
1.6.2 Evaluation the Framework.....	12
1.7 Thesis Contribution	12
1.7.1 Conceptual Contribution.....	12
1.7.2 Developing Contribution	13
1.8 Thesis Structure.....	14
1.9 Publication	15
Chapter 2: Literature Review	15
2.1 Medical Libraries	15
2.1.1 Overview.....	15
2.1.2 Type of Medical Library Science.....	16
2.1.3. Medical Libraries' Classification	17
2.1.4. Medical Library Services and Users.....	19
2.2 Physicians Needs and Information Seeking	21
2.3 Integrating Technology on Medical Libraries	26
2.4 Accessing Medical Information on the Internet.....	30

2.4.1 Information Retrieval using Search Engines	32
2.4.2 Information Retrieval using Mobile Phone	40
2.4.3 Information Retrieval using Online Resources	45
2.5 Factors and Barriers effect Seeking Medical information.....	50
2.6. Tools for Indexing Medical Information	54
2.6.1 MeSH.....	54
2.6.2 UMLS Unified Medical Language System	57
2.6.3 ICD-10 Codes for Information Retrieval.....	59
2.7 Important Issues Outlined the Proposed Solution	63
2.8 Summary.....	65
Chapter 3: Methodology and the Proposed Framework	66
3.1 Research Methodology	66
3.2 Scope and Objective Methodology	67
3.3 Framework Design Methodology	69
3.4 Introduction to the framework	71
3.5 Framework Architecture.....	74
3.6 Description of the Framework	75
3.6.1 First Layer: Access Layer	75
3.6.2 Second Layer: Control Layer.....	75
3.6.3 Third Layer: The smart Layer.....	78
3.6.4. Forth layer: Service Layer (Application Services)	78
3.7 Overview of the Protocols and the interfaces in the Framework.....	78
3.7.1. The Protocols	79
3.7.1.1 SIP Protocol	80
3.7.1.2 AAA Protocol.....	81
3.7.2. The Interfaces.....	82
3.7.2.1 Gm Interface.....	83
3.7.2.2 Cx Interface	83
3.7.2.3 Mw Interface	84
3.7.2.4 Isc Interface	85
3.7.2.5 The Sh Interface	85

3.8 User Profile	86
3.8.1 Overview of the User Profile	86
3.8.2 Public User Identity	88
3.8.3 Private User Identify.....	89
3.8.4 Initial Filter Criteria.....	91
3.8.5 Trigger Point	92
3.8.6 User Profile Data Request and Update.....	93
3.8.6.1 Data Requesting by the S-MICF	94
3.8.6.2 Data Requesting by the Application Server	94
3.8.6.3 User Profile Update	95
3.9 Framework Contraption.....	96
3.10 Summery	97
Chapter 4: Signalling Plan	98
4.1 SIP Addresses	99
4.2 Message Format	99
4.2.1 Start line of SIP Message Request and Response	100
4.2.2 Header Fields	102
4.2.3 Message Body.	105
4.3 The Registration	106
4.3.1 From the UE to the S-MICF	108
4.3.2 From the S-MICF to the MSDB.....	109
4.3.3 From the P-MICF to the UE	110
4.4 Authentication.....	111
4.4.1 Overview.....	111
4.4.2 Authentication of the Initial Register Request.	114
4.4.3 S–MICF Challenges the UE.	117
4.4.4 UE Response to the Challenge.	118
4.4.5 Successful Authentication.	119
4.5 Session Establishment.	120
4.5.1 Overview.....	120
4.5.2 Generating the Request.....	121

4.5.3 Processing the Request	121
4.5.4 Processing the Response.....	121
4.6 Example Scenario.....	124
4.6.1 First Scenario -General Search.....	124
4.6.2 Second Scenario Expert Search	126
4.7 Summary.....	129
Chapter 5: Smart layer Design Methodology.....	130
5.1 Introduction.....	130
5.2 Case Study	131
5.3 Design and Setup of the Experiment.....	133
5.3.1 The purpose of the Experiment	133
5.2.1.1 Selection of Queries.....	133
5.2.1.2 Web Search	134
5.2.1.3 Expert Search	135
5.2.1.4 Data Collection.....	135
5.4 Results of the Experiment.....	136
5.4.1 General search.....	136
5.4.2 Group 1 (Nurses)	139
5.4.3 Group 2 Dentists.....	141
5.4.4 Group 3 Pharmacists.....	143
5.4.5 Group 4 Medical Student.....	145
5.5 Discussion.....	147
5.6 Expert Search.....	149
5.7 Highlights of the Outcomes of the Experiment	152
5.8 Smart layer Description.....	154
5.8.1 The aim of the Smart Layer	155
5.8.2 The Design of the Smart Layer	155
5.8.3 Smart Layer Components	156
5.8.3.1 Query Adaptation Engine QAE.....	156
5.8.3.2 Results Adaptation Engine (RAE)	157
5.8.4 Smart Layer Search Parameter.....	158

5.8.5 Personalization and Ranking of the Search Results	160
5.8.5.1 Gathering the Information about the User	160
5.8.5.2 Learning Different Users' Interests	161
5.8.5.3 Ranking the Search Results Based on the User Profile	165
5.8.5 Handling Request over the System	167
5.9 Summery	168
Chapter 6: Framework Evaluation	169
6.1 Evaluation Methodology	169
6.1.1 Databases Performance Evaluation	169
6.1.2 OPNET	172
6.1.3 Jitter Delay	173
6.2 Case Study (User Classification)	175
6.3 Simulation Model.....	176
6.4 Results/Discussion	180
6.5 Summery	187
Chapter 7: Conclusions and Future Research	188
7.1 Conclusions	188
7.2 Achievements.....	189
7.3 Limitations	191
7.4 Future Work.....	193
References	196

List of Figures

Figure 2.1 Medical Information Retrieval tools	31
Figure 2. 2 Medsaerch Interface	35
Figure 2. 3 WRAPIN Interface	36
Figure 2. 4 Diagnostic Decision Tree.....	38
Figure 2. 5 Healthline Search Engine.....	39
Figure 2. 6 Ovid Interface	46
Figure 2. 7 MeSH Structure	55
Figure 2. 8 MeSH Database (MeSH 2016)	56
Figure 2. 9 MeSH Training (MeSH 2016)	75
Figure 2. 10 ICD 10 (WHO 2015)	61
Figure 3. 1 Research Methodology General Flow	67
Figure 3. 2 Organizational Diagram of the Framework Design	69
Figure 3. 3 IMS Architecture	73
Figure 3. 4 MICF Architecture	74
Figure 3. 5 The MSDB Interfaces.....	77
Figure 3. 6 IMS Reference points.....	79
Figure 3. 7 Overview of Protocols and Interfaces	79
Figure 3. 8 User Profile	87
Figure 3. 9 Public and Private User Identities	91
Figure 3. 10 Initial Filter Criteria	91
Figure 3. 11 Server Trigger Point Structure.....	93
Figure 3. 12 Process of S-MICF Data Request from MSDB.....	94
Figure 3. 13 Process of Application Server Data Request to MSDB.....	95
Figure 3. 14 Process of User Profile Updates	95

Figure 3.15 Data request by the MSDB	96
Figure 4. 1 SIP Message Format.....	99
Figure 4. 2 Header Field for SIP REGISTER Message	104
Figure 4. 1 Header Field for SIP 200 OK	105
Figure 4. 2 Registration Procedure	108
Figure 4. 3 Register header field.....	110
Figure 4. 4 SIP 200 OK Field	111
Figure 4. 5 IMS Authentication	112
Figure 4. 6 Call Flow for the Authentication	115
Figure 4. 7 Call flfor the INVITE Message	122
Figure 4. 8 Header Fields for the INVITE Message.....	123
Figure 4. 9 SDP for 200 OK	123
Figure 4. 10 Request Data Flow Diagram Scenario 1	125
Figure 4. 11 Request Data Flow Diagram Scenario 2.....	127
Figure 5. 1 Organisational Diagram of the SL Design.....	130
Figure 5. 2 Search Results Obtained by Nurses and Medical Students for Diabetes	132
Figure 5. 3 The Three Search Queries.....	134
Figure 5. 12 General Search Results on Google	136
Figure 5. 13 Results Obtained by the General User	138
Figure 5. 14 Search Results Obtained by Nurses	141
Figure 5. 7 Search Results Obtained by Dentists	143
Figure 5. 8 Search Results Obtained by Pharmacists	145
Figure 5. 15 Search Results Obtained by Medical Students.....	147
Figure 5. 16 Expert Search Results	152
Figure 5. 17 Search Results for Children and Diabetes	162
Figure 5. 18 Users' Profile Interest	163
Figure 6. 1 Abstract Model	171

Figure 6.2 Steady Stream of Packets	173
Figure 6.3 Handling Jitter	174
Figure 6.4 Handling Excessive Jitter.....	174
Figure 6. 5 OPNET Simulation Model.....	178
Figure 6. 6 End to End Delay CDF	181
Figure 6. 7 Voice Traffic Delay Variation CDF	182
Figure 6. 8 HTTP Request Throughput CDF.....	183
Figure 6. 9 Scenario 1 No Traffic Filtered.....	184
Figure 6.10 Scenario 2 50% Traffic Filtered	185
Figure 6.11 Scenario 3 75% Traffic Filtered	186
Figure 6.12 HTTP Requests Processing Delay.....	186

List of Tables

Table 2. 1 Medical Databases	48
Table 3. 1 SIP Request Methods.....	80
Table 3. 2 SIP f Response Codes.....	81
Table 3. 3 Cx Interface Command Name and Code.....	84
Table 3. 4 Commands Defined by the Diameter Application for the Sh Interface	86
Table 4. 5 SIP Request Methods.....	101
Table 4. 6 SIP Response Code.....	102
Table 4. 7 SDP Type and Value	106
Table 5. 8 Number of General Search Results by Search Method	137
Table 5. 9 Search Results for the Nurses	140
Table 5. 10 Search Results for the Dentists	142
Table 5. 11 Search Results for the Pharmacists	144
Table 5. 12 Search Results for the Medical Students.....	146
Table 5. 6 Expert Search Results	150
Table 5. 7 Expected Results for the Search Topics	164

List of Abbreviations

AAA	Authentication, Authorization and Accounting
API	Application Programming Interface
AS	Application Server
Av	Authentication Vectors
CDF	Cumulative Distribution Function
CF	Collaborative filtering
CSCF	Call Session Control Function
HSS	Home Subscriber Server
HTML	Hypertext Markup Language
IMS	IP Multimedia Subsystem
IR	Information Retrieval
MIR	Medical Information Retrieval
ML	Medical Libraries
PC	Personal Computer
PDA	Personal Digital Assistant
PSCF	Proxy Session Control Function
MAA	Multimedia-Authentication-Answer
MAR	Multimedia-Authentication-Request
Mesh	Medical Subject Heading
MICF	Medical information Control Function
MIRS	Medical Information Retrieval Systems
P-MICF	Proxy Medical information Control Function
QAE	Query Adaptation Engine
RAE	Results Adaptation Engine

RS	Recommender System
SAA	Server-Assignment-Answer
SAR	Server-Assignment-Request
SDP	Session Description Protocol
SIP	Session Initiation Protocol
S-MICF	Server Medical information Control Function
SL	Smart Layer
UE	User Equipment
URI	Uniform Resource Identifier

Chapter 1

1.1 Introduction

Due to rapid developments in the health care domain, the need for extensive information has become one of the urgent requirements demanded by physicians around the world. Also, with the development of information technology and the Internet, it has become necessary for medical libraries to face emerging challenges and invest in these technologies in the right way to provide better services to their users. Therefore, online medical libraries have tried to cover the large scale of distributed knowledge across the web by using the latest developments in technology through building systems that provide high speed and flexible retrieval (Wenyun et al. 2010). These functions are beneficial for both healthcare professionals and for the medical libraries themselves. Among the most important methods that physicians use to seek information, this research finds that medical libraries are one of the top-rated search methods. Furthermore, physicians also access medical resources from various devices. According to Dee et al. (2005), 87% of physicians have used a PDA for locating information relating to patient encounters. Other methods include browsing the Internet and accessing information through three main mechanisms: search engines, directories and subject-orientated websites, and each of these search tools yields a different set of results (Wukovitz, 2001). Online library catalogues can also be used as a way to access medical library resources, including databases, e-books, journals and many more. Despite all the methods mentioned above, locating information has become increasingly complicated because of the vast amount of information in the digital environment and not just on health domain. Medical information websites have solved

some problems of searching health information; however, there are still other problems that physicians face, such as the need to register on some sites and the need to record data, in addition to other difficulties such as learning how to access catalogues or learning resources (Wukovitz, 2001).

To change this situation and solve these problems, efforts have been made to improve different methods for searching for health information. These efforts have mainly focused on designing specific medical search engines, such as Helpfulmed (Chen et al. 2003) and MedSearch (Luo et al. 2008). Others have focused on designing a retrieval system that can be used in medical libraries, such as the centralised system for personal information retrieval, and access services made by Che and Dahanayake (2004). Other approaches have been developed by groups of researchers as a solution for obtaining specific information. These have focused on developing innovative technologies and scenario-based proxies that enable the gathering and filtering of information customised for users. This approach might be helpful to provide relative medical information for users (Chu et al. 2000). Other study introduced a retrieval model based on web services to improve the retrieval system in digital libraries (Liu et al. 2010). Another related study designed a digital library information retrieval model based on ontology (Wu-Jun and Peng 2010).

Mobile phones also pose something of a challenge since smartphones have appeared as a tool for accessing information. A number of studies cover this subject and several frameworks have been designed to facilitate access to online resources from a variety of smartphones. An example of these frameworks are Portal Personal Spaces (Castellanos and Sanchez 2003), and the Service-Oriented Medical Framework (Park and Nam 2009). Furthermore, medical libraries themselves have started to provide user services and build accessible applications via mobile devices. For example, SMS services were provided by the international Islamic University in Malaysia (Abdul-Karim et al. 2006), and library

catalogue searching at Delta State University (Iwhiwhu, et al. 2010). Unfortunately, the attempts made on mobile library access were limited on some services, including SMS and library catalogue search services. Despite the successes of some attempts made, others revealed a number of problems such as the limitation on specific services, providing general searches for the huge number of distributed databases on the Internet, lack of structure on mobile applications that were limited on SMS services and accessing library catalogues and others that will be discussed in the second section of this thesis. Therefore, this research aims to remedy the problems identified and overcome the limitations to meet users' needs by providing and introducing a universal framework that aims to personalise the search result based on the user profile.

1.2 Research Problem and Motivation

Medical information retrieval systems (MIRS) refer to the methods and technologies that designed to assist the medical information seekers to access their information at the right time and the right place. Chapter 2 covers variety numbers of medical information retrieval methods which aimed at providing information to the physicians based on their needs. These methods arose a several number of issues that contributed significantly in highlighting the problem and confirmed of the needs of our solution.

There are several reasons why there is a need to design the framework and provide personalized medical information retrieval to understand the user' needs and customize search results based of the learned profile for the individual user. Apart from the steady growing rate of the medical researches the physicians themselves have faced a number of barriers while accessing the medical information from the existing systems. These barriers have been discussed in chapter 2. Moreover, the literature has identified the following issues:

1- Large amount of medical literature

Most of the studies have acknowledged of the rapid growth of medical literature. This growth is a result of the increasing number of research to investigate about the disease, treatment and technologies. This growth of the medical researches lead to increasing the number of published studies on the web which became hard to find a specific information.

2- Limited in scope

Some of the current methods were limited in scope and subject such as the efforts made on medical mobile access. Moreover, some system designed to be used by the physicians to request information regarding the patients' medical information, while physicians still need to have access the service out of the working hours.

3- Physicians needs

Designing a medical retrieval system needs to focus on the user needs but in number of cases the physicians themselves are unaware of their information needs. According to Gonzales (2007) physicians claim to the needs of daily information to answer patients questions, and because of the information requested by patients every day and other information they need for their knowledge it became even hard for them to be aware of their needs.

4- Time consuming

Due to the limited time of surgeons in general and during patient's consultants in particular retrieving medical information is complicated especially with the huge amount of information stored in the web. Although of that patients still demand to answer their question. According to Del, et all (2014) physicians frequently receive questions about patient care in their practice about 34% of these questions related to drugs and treatment, and 24% connected to causes and symptoms, and because

of the clinicians' lack of time some of these information may not answer the patients' questions.

5- Insufficient information

Due to the vast amount of medical information and to the lack of awareness of clinicians needs and due to clinician's limited time, the demand for further information is still requested by physicians as result of the insufficient information they retrieved in order to get answer for many questions.

6- Lack of experience and knowledge

Experience in performing the medical search and using verity of medical resources and selecting the suitable medical terminology play a significant role on retrieving a relevant information. However, limit knowledge and lack of experience on performing the search will effect receiving relevant information.

In addition to the above issues, this study presents important issues that outline our solution presented in chapter 2 section (2.7)

1.3 Research Motivation

The researcher' background is in library and information science and the area of information retrieval, health care and medical information retrieval system is a new involvement for the researcher.

The need for this research is motivated by two issues mentioned below. Thesis issue are dealt with in depth in course of the research and expected out comes are under the heading, objectives

1- Users' needs

Medical information community contains of variety of users from different medical specialties such as surgeons, paediatricians, pathologist and many more, each one of them has different needs based on their area of interest and sure there will be some common needs related to diagnosis, treatment and patients 'advice. However, understanding individual user' needs is one of the issues that been disused in a number of studies in chapter 2. The attempts made previously have taken into consideration the user' needs and have designed a number of retrieval systems for professional from medical domain and the other approach has even considered the patients and their limit knowledge in medical therefor they have designed medical search engines that have ability to understand the simple user queries.

Most of the development in the medical information retrieval focused on the relevance and how to provide the user with relative information and this is the aim of any information retrieval system. However, personalizing the medical search results is another area of challenge in the IR, this challenge did not discuss in the previous attempts. Therefore, this study aims to address the user needs and customizing the search results and that is a greatest challenge in this study.

2- Users' knowledge and skills

As previously mentioned that medical information seekers divided into number of categories based on their major. Each group of users of medical information retrieval systems has different levels of medical knowledge, this level of knowledge is also different for the individual user within the same category. The variety of knowledge among the users will definitely impact the way in which individuals perform the search queries to the systems. And that is also has an impact of the search results and the way that the system returned the results to each user. Therefore, this study has taken into

consideration the user's knowledge. The variety of knowledge has contributed to divide the beneficiaries into two categories in our study first is the general user including the patients and second is the expert user who is professional user from medical community who have more knowledge on performing a complex medical queries. This division showed the role for which the smart layer is designed.

1.4 Research Objective

The primary objective of this research aims to introduce a generic framework that allows users to access medical information from a wide range of devices. This should enable the medical professional to have flexible access to the medical information and resources in their everyday environment in addition to access to the latest information in the medical domain. Moreover, this research aims to introduce the smart layer that aims to achieve the goal of customizing the search results to get the most relevant information based on the user profile.

The framework is designed based on a layered architecture similar to that used in the communications system for the design of the IP Multimedia Subsystem (IMS), the layered architecture has been adapted to perform the tasks in the proposed system.

The need for designing the framework is achieved by the following objectives:

- 1- Study the existing research and published literature regarding accessing medical related- information and identify different type of methods used by medical professional to access their information.
- 2- Identify the problems from the existing solutions and determine the pros and cons of each study.

- 3- Identify users' needs from medical libraries by defining user groups and determining their needs. This can be achieved by designing a survey that will be completed by members of the healthcare community.
- 4- Study the communication system IMS architecture.
- 5- Introduce and present a comprehensive framework, describe the architectural layers and define the main functions of each layer including the connectivity (protocols and interfaces) and signalling plan for the message.
- 6- Identify the needs and the functions for the smart layer and design the smart layer.
- 7- Determine the appropriate tool for the evaluation, and the evaluation performance metrics to validate the system performance and to generate the results.
- 8- Generate the results and draw the evaluation conclusion.

1.5 Scope of the Study

Information retrieval in general, is a huge topic studied by number of researchers. The field of medical information retrieval in particular has gained a significant interest from the researchers after 20 years of the invention of information retrieval itself in 1940 Can and Baykal(2007).

Since then a number of studies have conducted in medical information retrieval which can be divided into two main aspects first one is focused on the stand -alone retrieval systems that used inside the medical institution which includes laboratory results and assessment systems , patients related information records system and images and x-ray retrieval system. The second aspect is the retrieval system used by physicians to access medical information and to enrich their information on their daily life which includes medical databases and web medical browser. This research will not focus on the first aspect because they are related to patients

related- information which can be used in the hospital environment and this is out of the scope of this thesis. Therefore this research will study and investigate the current medical retrieval system and identify the pros and cons of them that led to the main problem of this study and focus on the following.

- The thesis focus on designing a generic framework for accessing medical information to solve the problems of retrieving large amounts of medical information, and avoid the limitations that arose in the previous attempts discussed in chapter 2 (literature review).
- The study will not focus on improving the retrieval system in general, but it will focus on a special case which is medical information and the relationship between the system and personalization of the system for medical information which is managed by the smart layer. Therefore, it will focus on the user behavioural on how certain type of users from medical domain use the information retrieval system for medical purposes. And then to approve the role of the smart layer that aims to achieve on customising the search results based on user's profile.
- The focus of this study will be mainly on the main control unit which holds the main unit Medical Information Control Function MICF and the smart layer. This study will not focus on providing the services to the user through application program interface API because that is a different area of which this study aims to. Also it will not focus on the application layer because this is out of the scope.
- The system will be validated using a simulation model (OPNET) to estimate the system performance measures such as the delay and throughput under a heavy load of request by the users.

In brief, the focus of this thesis has two aspects, first is technical and that focuses on the designing of the framework which will be covered in chapter 3, and the second is functional which focuses on presenting a call flow process for the way that the system handle the transaction and that is covered in chapter 4.

1.6 Research Methodology

The methodology that has been used in conducting this research is summarized in this section and a detailed description of the framework methodology will be presented in chapter 3, and the methodology of designing the smart layer will be presented in chapter 5. The study methodology is divided as follows:

1.6.1 Designing the Framework

The methodology of the research followed in this research is based on the methodology found in the literature such as the health monitoring system by Rikitake et al (2009), Mobile Telemedicine System by Navarro et al (2006), mobile-Health Care System by Mayuri Gund et al (2011), and Medical Sensor Application Framework by Markota & Ivica (2010).

The framework is designed based on layered architecture similar to that used in the communications domain for the design of the IP Multimedia Subsystem (IMS) the layered architecture has been adapted to perform the tasks in the proposed system. The framework adapts the layered module from the IMS and presents the following four main layers **access layer** which adapts different entry device, **control layer** that controls and manages the requests and holds the main control unit which called Medical Information Control Function (MICF), **smart layer** which designed to customize the search results and supports the user request based on their profile and finally **the application layer** which is basically set of medical library

services. Each layer contains number of units which are aim to perform number of functions. The details of the framework entities are presented in chapter 3.

1.6.2 Framework Evaluation

The framework evaluated and validated using a simulator model (OPNET). A simulation model of a new framework for digital medical libraries along with a benchmarking technique are introduced to best evaluate the performance of the medical database library and to estimate the system performance measures such as the delay and the throughput under heavy load of user requests. More details will be covered in chapter 6.

1.7 Thesis Contribution

The thesis details the contributions made to knowledge in two ways- the conceptual contributions and the development contributions.

1.7.1 Conceptual Contribution

Understanding and investigation of the existing medical information retrieval systems used by physicians to seek their information which led to the following contribution:

- Several studies on medical information retrieval system do acknowledge of the main role of the medical information retrieval systems play on providing relative medical information to the physicians whether it was used during the patient's consultant or in their daily life. This study has contributed to determine and classify the methods used by medical information seekers to access the medical information. There are clearly three main methods offered to the medical information seekers, first category is medical information databases, second category medical information websites and

lastly medical information mobile applications these methods are presented in chapter 2.

- The second conceptual contribution of this thesis is a technical perspective of the medical information system and that is presented as a layered model architecture. The details of architecture will describe in chapter 3
- While most of the existing medical information retrieval system address the problem of providing the physicians with medical information retrieval system to support them during the patient's consultants and assist them on making decision and treatment plan. Although some of them address the user's needs but did not focus on how to customize the search result in order to avoid the large amount of the returned results. As a result of that this study has introduced the smart layer as a third contraption of this thesis. The smart layer will present in chapter 3.

1.7.2 Developing contribution

The developing contribution arise from the evaluation of the framework which presented in chapter 3 and chapter 4.

- The primary contribution of the developing contribution is designing of the framework and setting the protocols and interface for connection.
- The second developing contribution is providing a details of the registration process, authentication process and handling the session over the system.

The designing of the smart layer is another area of contraption that presented in chapter 5 to test the validity of the framework and to approve of the need of the smart layer for customizing the search results based on user profile.

1.8 Thesis Structure

The thesis is structured in a way to let the reader understand the problem, gain general background information about the topic, have an idea about the existing medical information access methods, have clear idea about the problems that physicians faced while accessing the current systems and finally to understand the proposed framework and evaluation methodology and results. In more detailed the thesis is organized as follows

Chapter 1: Introduction

An overview of the scope and the concepts leading to the study, along with the objectives that the study aims to achieve is presented in the current chapter. The research problem is established in this chapter and will be detailed in chapter 2.

Chapter 2: Literature Review

An introduction to medical libraries and their types, users, and services are illustrated in this chapter. It then discusses the medical information retrieval methods and provides a variety of studies that done on accessing medical information using medical search engines, medical online resources and mobile based medical application. The chapter then reviews a number of existing medical retrieval (related systems) for accessing medical information.

Chapter 3: Methodology and the Proposed Framework

This chapter introduces and presents a comprehensive framework, and gives a detailed description of the main layers of the framework. The main component in each layer and their functionalities together with the protocols and interfaces linked them with each other are detailed.

Chapter 4: Signalling Plan

Signalling plan and call flow of the process will be illustrated. This chapter covers the message format and then gives a detailed description on handling the message and session setup. It will then present two scenarios of two types of search (general search, expert search).

Chapter5: Smart Layer Design Methodology

This chapter introduces the smart layer design methodology and the steps that led to the need of the smart layer to be introduced in our framework.

Chapter 6: Framework Evaluation

In this chapter the designed framework will be validated and evaluated via a simulation model and presents the result of the evaluation. The evaluation will estimate the performance measures such as the delay and throughput under heavy loads.

1.9 Publication

Ali, Ashraf & Ware, Andrew & Alzahrani, Fatimah & Al-Begain, Khalid. 2018 "Performance Evaluation and Benchmarking of Mission Critical Medical Access Systems," *2018 IEEE 20th International Conference on e-Health Networking, Applications and Services (Healthcom)*, Ostrava, 2018, pp. 1-6.

Chapter 2

2. Literature Review

2.1 Medical Libraries

2.1.1 Overview

When referring to libraries in a health environment, means that referring to the healthcare and biomedical information, research, sources and services. A medical library is basically a large collection of the above mentioned services that is not different from other libraries in that it has the same design of services (Catalogue, references, index, and more) managed by medical organization, but with a different collection (medical collection). It can be accessed physically or through the Internet.

Medical libraries started for decades. In ancient Assyria, China and Egypt there were collections of many medical books. For example Egyptian medical papyri are dated back to between 1900 and 1200 B.C. One library of an ancient king of Assyria called Ashurbanipal was written between 665 and 626 BC. This particular library had more than thirty thousand fragments of clay tablets, 800 of them being medical. Of all the libraries of the ancient world those of Alexandria - the Serapeum and the Bruchium – were known of in most of the regions at that time. Most studies of the Romans depict that they did not carry medicine with high regard. However the Byzantium libraries had many works on medicine (Chen et al. 2011). Shah (2011) notes that between the 9th and

the 11th century when Islam reigned supreme in science there were many medical libraries that were built in Spain and in the East. Before his death, Avicenna had comprehensively described the Royal Library at Bokhara. During the 3rd century the founders of monastic libraries provided for the control and custody of books. When students began to be taught medicine in the universities that were later founded medical libraries were built but their growth was rather slow. However at the same time the medical libraries of many societies and corporations had already been established. A good example is the Royal Society of Medicine (RSM). Up to date the RSM is the leading medical library in Britain (Connor, 2011). The RSM has grown to become a charitable organization that provides medical education all over the world. Every single year it holds more than 350 conferences and meetings. RSM publishes the Journal of the Royal Society of Medicine in addition to providing vast information on those students who want to pursue careers in medicine. In 1836 the National Library of Medicine (NLM) was founded in the United States and it has become the largest and most useful medical library all over the globe. It is located in Maryland and among its collections are some close to eight million journals, books, manuscripts, technical reports, images and photographs, and microfilms on medicine and other related sciences. The library also facilitates the National Center of Biotechnology Information (NCBI) (Boruff and Bilodeau, 2012).

2.1.2 Type of Medical Library Science

In the medical field the libraries can be divided into three main types: academic, hospital and special. Each library has its own features but shares the same functions

- 1- Academic health science libraries that located in a medical educational institution to the faculty and the students. It contains a number of information

resources and learning materials along with the scholarly resources used by the academic staff for teaching purposes.

- 2- Hospital libraries: allocated at the medical centre and contains a large collection of resources and services. the primary purpose of the hospital libraries is patient care but they also engage in preventive care and health education ,community service, professional research
- 3- Special libraries that include all type of medical science libraries not mainly related with an educational or healthcare insulation.it contains all information associated to health. There are many number of Special libraries including those serving the army and the federal health agencies such as the (national library of medicine and national institutes of health). Or it can be a professional societies and organization such as the American Hospital Association and the American College of Obstetrics and Gynecology (Sandra 2014)

2.1.3 Medical Libraries' Classification

Historically, libraries have been classified as national, public, academic, special or school libraries (Connor 2011). In this case, one may consider medical libraries to be special libraries because the services they offer and their clientele are geared towards a certain subject and community. However, they are more than special libraries because they are academic, national and public libraries as well. In addition, the information in these libraries doubles every 15 years or less, the number of students enrolling in medical schools has also increased immensely (Connor 2011). The nature of the medical field has also been evolving diversely, as medical literature has begun to take into account the sophisticated interplay of emotional, environmental and personal factors that impact the health of people. Also, there has been increasing interest in fields

like sociology, economics, psychology, anthropology, operations research, communication science, electronics, mathematics, ecology and biophysics (Boruff and Bilodeau 2012).

Generally, there are various ways of classifying medical libraries, from the type of user they serve, the subjects covered (e.g., drugs, x-rays, etc.), and the material collected. For example, a nursing library may be covering nursing as a subject, serving the nursing community and collecting various materials for researchers. The most common mode of classification is outlining the parent organisation, which may be either public or private (Encanto 2011).

- First, there are national medical libraries. In a national medical library information is collected and achieved regardless of where it comes from. One example is the United States National Library of Medicine, which serves people all over the world with its virtual information. Another example is the National Medical Library of UAE University, which also has information on the heritage of the UAE (Kapronczay et al. 2011).
- Second, there are medical school and medical centre libraries. this category includes medical libraries of institutions that teach medicine, pharmacology and nursing. These kinds of libraries vary vastly in size, staff, space, service provision and automation (Plutchak 2012).
- Third, there are hospital libraries. These are the libraries that have been charged with the responsibility of making sure that the clinical, administrative, research and educational personnel have access to information that is crucial to the provision of adequate patient care. Such libraries also vary depending on size and other variables. In

some cases, hospital libraries are services that are provided by public libraries (Kapronczay et al. 2011).

- Fourth, there are the libraries of other health care entities. These may be private or public entities. In addition to the resources they have internally, they have access to more information from the Internet and other resources. Other medical libraries include government agency libraries, libraries of voluntary health organisations, professional societies, pharmaceutical and other company libraries (Kapronczay et al. 2011).

2.1.4 Medical library services and users

There are many services that are provided in a medical library. Most of the online medical libraries also offer access to a number of sources and services such as access to e-books, e-journals, e-magazines, research papers and educational software, among others. MLs also offer research and curriculum support to enable users to consult with librarians over the Internet. Users can also access online consultation services including reference services, educational support, bibliographic management, connectivity issues and remote access, networking for hospitals, clinics and medical students, presentation skills, liaison and outreach activities, information on consumer health resources, library publications and alternative health resources, multimedia resources, blackboard and other services on curriculum support (Plutchak, 2012). Many libraries offer various computer-assisted instruction resources and software. These include Blackboard, Springer Images and medical education software in anatomy, cell biology and histology, cardiovascular medicine, microbiology, neuroscience, nursing, clinical skills, pathology and laboratory medicine and obstetrics and gynaecology (McGowan 2012). Some of these services required subscription and some of them are available to the user for a certain duration of time.

Medical libraries also offer education to users on how they can publish their work in journals. This is usually made for publishing research among students and other scholars. The services provided covered how to publish, conduct research, write and submit papers. Institution-based medical libraries also offered students the option of having a personal librarian. This is usually a programme that is meant to assist students in medical school during their years of study. This programme helps to find a librarian for each student, who assists them with library-related questions (Shah 2011). The online personal librarian recommends resources to the student in a manner as to meet his or her personal needs. The personal librarian also instructs the student on new resources and technologies in addition to assisting them with accessing particular resources as their learning needs change with time. Many medical libraries also offer photocopying and printing services. The users are charged depending on the number of pages they want to print or copies they want to make.

The mentioned services are available to a number of medical users, those users can be divided into two categories: organisations and individuals. Organisations can be governments, research institutes, schools, colleges and universities, business corporations, hospitals and other organisations. Individual users can be student researchers, the general public, health professionals such as nurses, surgeons, and medical doctors, and allied health workers like laboratory technicians (McGowan 2012).

2.2 Physicians needs and Information Seeking

Physicians are constantly in need for relevant and new information pertaining to patient care, drugs as well as newer forms of diseases.

Generally speaking, Physicians can base their decision for patient's consultants and treatment plan to a tow types of knowledge:

- 1- Explicit knowledge, related to the already well established and formalised domain knowledge such as books and other textual documentation or can be represented by health professional and developer such as physicians' guideline.
- 2- Implicit or tacit knowledge, this found in the individual expertise, medical organisations practices, and past cases (Montani and Bellazzi, 2002).

According to Altschuler et al. (2012), a physician would need to spend at least 21 hours a day to provide high quality care for patients, including acute, chronic and preventive health care treatments. As a result of high patient numbers and fewer doctors, the average patient consultation time has reduced to about seven minutes per patient, which does not give the physician much time to search for information pertaining to patient's specific questions and needs (Altschuler et al., 2012; Gonzalez-Gonzalez et al., 2007). On average, a patient asks doctors between one to two questions per consultation session, which usually revolves around the diagnosis and the treatment options (Gonzalez-Gonzalez et al., 2007). It was reported that close to twenty-three percent of physicians tend to search for answers using various medical information tools for possible questions, including during consultation times.

The need for quick and relevant answers pertaining to diagnosis, symptoms and possible treatment options, are essential to save time during patient and doctor consultation. It was also reported that the average time it takes to find a successful and relevant answer was around two minutes per question during consultation time, which could be improved to accommodate more questions from concerned patients, as well as to improve the quality of the consultation (Gonzalez-Gonzalez et al., 2007).

According to Kannampallil et al (2013), the human mind has a limitation in terms of storing and recollecting memory, as well as being unable to cope with an overload of information. These limitations often affect the ability for physicians to make optimal decisions, and therefore, there is a need for accessing information to boost the memory and to make the best informed decision possible, since there are a number of medical databases such as Medline and PubMed and many more that provide information availability to the user that assist the medical decision making process (Johnson, 2008.; Luo et al 2008., Leo et al., 2006)

However, in order to get to the right information that meet the physicians' needs, users should be highly skilled in selecting the right terminology and performing their search using different search methods and this is what the information seeking behaviour is aimed to. Information seeking behaviour is important aspect of physicians, nurses and those in the medical field as information plays an important role in determining the possible decisions and options to diagnose and to treatment and care for patients (Reddy and Spence, 2008). Moreover, the information seeking behaviour also helps to assist in the decision making process, which requires physicians especially to be interested and prepared to seek for the relevant information. The retrieval of information in the medical field requires the seeking behaviour that goes through keywords, terms as well as different sources to identify the right option.

The collaboration between physicians is also an important part of the seeking behaviour when it comes to the need for information. Reddy & Spence (2008) reported that collaboration is a normal part of research, particularly in the medical field, and the information seeking behaviour of teams may differ compared to individuals due to the different searching methods and abilities. Their study focused on identifying the need

for information in teams and the situations in which the need for information triggers a collaborative effort to find results together. A total of seven needs were identified from the study. It was found that the majority of the needs were related to finding specific information for patients, and how to care and treat patients.

Another point that reported by Reddy & Spence (2008) for the information seeking behaviour is the needs of collaboration efforts when there was a lack of expertise among physicians and nurses who were treating patients together. Since there is insufficient information and doubt arises, there is a need to collaborate and find information immediately to provide options for the patient (Reddy & Spence 2008). Furthermore, collaboration also took place when there was a lack of available and accessible information. This causes team effort to communicate and to identify the best possible option through existing information. Lastly, collaboration also took place when there were challenging information needs, such as new diagnosis of disease and rare conditions. Hence, it is vital for physicians and other medical team members to communicate on a personal level in order to collaborate when seeking for information, and that this will cause the information seeking behaviour to yield in better results (Reddy & Spence 2008).

The medical field is viewed as an evidence-based field, whereby physicians need to consistently keep abreast with the latest findings and information pertaining to their expertise. An average physician needs to use at least two million different types of information in order to manage their patients (Naidoo et al 2010). The newer developments in the field is how the medical field grows, and for this to happen, doctors have to be constantly accessing to the information and then assessing and analysing data so that the patients are able to benefit from the care. However, time constraints

often times deter physicians from being able to access different information, especially due to the high number of patients in their care. It is also almost impossible for physicians to know everything about their field since the information that is available is just too vast, especially over the internet (Ajuwon & Popoola 2014). Furthermore, in cases where patients require all options, particularly for surgery cases, a physician needs to be well-equipped to provide answers to the patients in detail. Having useful relevant and recent information can help to build trust and relationship between patients and physicians. Moreover, physicians also need to look up the latest pharmacological information in order to prescribe newer and updated drugs for their patients. Some physicians also search for information merely out of curiosity and personal interests in updating their knowledge. Evidence is another reason why doctors' search for information (Ajuwon & Popoola 2014).

It is important that the information they get from the internet has to be relevant, informative and organized as this will help them to use the information and at the same time, save time from having to view multiple websites or search engines (Ajuwon & Popoola 2014).

Unfortunately, often times, the information is often provided to physicians in an unorganized manner, whereby the physician has to look at different types of existing information, before finding the information that is required in that situation in order to make informed decisions (Kannampallil et al. 2013). Furthermore, some of these information are stored on online medical records, which once again requires the task of finding for the specific patient file, sorting and retrieving information as well as filtering

out unwanted information to gain specifically required information in order to make informed decisions for patients.

Undoubtedly, that the availability of medical information recourses have provided positive outcomes for physicians on a daily basis, especially since there are constant changes and improvements made in the medical field (Gonzalez-Gonzalez et al. 2007). According to Kannampallil et al (2013), the need for information is vital for physicians as they have to find relevant information and apply it to a given situation. The lack of information or the use of outdated drugs and information could lead to negative outcomes and perceptions for patients. Hence, there is a need for physicians to be equipped with up to date and relevant information that could be accessed instantly. Although the availability of information is widespread through different sources such as through electronic records, the need to search, filter and manage useful information still exists (Kannampallil et al. 2013).

Besides that, doctors could also find it difficult to provide high quality care since the decision made can have a positive or negative outcome for the patient, depending on the available information (Kannampallil et al. 2013).

Physicians' needs from the medical library are diverse based on their specialities, knowledge, experience, interests and other needs. Davies (2007) highlighted the physicians' needs and divided them into two types: first is primarily perceived and second is unperceived needs. According to perceived needs, the need for information is recognized and attempts are made to retrieve or find an answer. On the other hand, unperceived needs are not categorized as a type of information that should be addressed. This creates a gap in the physicians' existing knowledge in the medical field. However,

Bukachi and Pakenham-Walsh (2007) disagree with the fact of dividing the types of need for information, since in most cases the information is provided only when it is required, and therefore this does not create actual need.

Ajuwon & Popoola (2014) on the other hand, identified seven factors that affect the need for physicians to search for relevant information. The first need is pertaining to patient consultation and care, whereby the information is passed on to the patient for their benefit. The second reason focuses on drugs and treatment, the third and the fourth are related to accessing resources and information to assist them in the decision making process, particularly if it is a risky or difficult option for the patient. The fifth and sixth is related to medical training and practical workshops while the last reason for physicians to search for information is to keep abreast with the latest clinical findings in their field, and to reduce the gap in knowledge (Ajuwon & Popoola 2014).

2.3 Integrating Technology on Medical Libraries

Since the appearance of the electronic resources, the librarian have moved their focuses toward providing digital library service and replace the traditional services with an electronic services an example of that is the references services that were provided to the users based on a conversation between the librarian and the user to determine the user needs, then the reference librarian is responsible of collecting the references and picking up the most appropriate sources for the user at the given point of time. This process has changed within the digital library services, users became able to select the appropriate source on digital environment. And nowadays the electronic library services playing a main role on providing a variety of online form of current awareness and selective dissemination of information services, which aim to inform the user of the latest information in their areas of interest (Chowdhury, 2002). And with the

increasing number of the electronic resource and the emergence of the web and libraries. It became necessary for the libraries in general to invest the technology in different area to provide better services to the user.

The development of information technology also has made it necessary for medical libraries to keep up with changing trends and times, and to focus on integrating these technologies in the most efficient way that could generate better search results and retrieval methods for medical users. According to McGowan (2012) in his study reported that the future of medical libraries are definitely heading towards online materials. The mixed method study employed the use of questionnaires and interview method to analyse the future of medical libraries. The focus of the participants were directors of academic health sciences libraries. The study concluded that almost a hundred percent of participants had purchased online journal articles, more than ninety percent of participants were signed up for subscriptions with medical databases, and more than quarter of the total respondents had purchased e-books online compared to reading hard copy books. Besides that, the study also found that close to twenty percent of participants indicated that this enabled them to save time, and possibly continue to save more time in the future due to improvements made in purchasing options online (McGowan 2012).

Most of the medical libraries also have incorporated technology as part of training for health professionals in order to encourage adoption of technology, and to focus on personal development of medical professionals (Graham-Jones et al 2012). Information technology has become a part of the medical field and it is inevitable, that medical professionals need to be well aware and comfortable to use technology in a way that will help to boost and improve productivity, particularly when it comes to retrieving important medical information from the online medical libraries (Graham-

Jones et al 2012). It was estimated that it would cost more than twenty billion dollars for medical libraries to embrace and integrate technological systems, however, the benefit of this integration has been viewed as highly beneficial for the users (Graham-Jones et al 2012). One of the suggested ways of integrating technology into the medical libraries is through the education stage, whereby medical trainees are trained and expected to use medical databases to access and gain information pertaining to patient's diagnosis and treatment options (Graham-Jones et al 2012). This is expected to familiarize medical students with the use of technology, and at the same time train them the importance and the usefulness of the system, so that it will be practiced during consultations and diagnosis (Graham-Jones et al 2012).

According to Graham-Jones et al (2012) reported that there are implementations to include technological information retrieval services that provided by the medical libraries in medical organization, the rate of adoption is relatively high, which means that there is a high impact of the use to materialize for increasing patient care and quality. That was confirmed by third of medical information seekers who indicated to the importance and usefulness of the library resources in the medical field (Graham-Jones et al 2012). Although that there are some findings and results that have proven the lack of awareness among the physicians and medical students of using the library resources and services but most of them have confirmed of the role that the medical library plays on training the users on how to use the services and also have confirmed of the need of the ML resources and services.

Medical libraries also have integrated another methods to increase the usage of technological devices and applications to retrieve information an example of these methods is using personal digital assistants (PDA). PDA is a hand held tool which

used by the health care professional the device is loaded with suitable functions and software applications that allow the clinicians to retrieve medical or personal information (Lindquist et al., 2008). The usefulness of using the PDA to retrieve medical information is discussed in a number of studies such as (Mechling et al., 2009);(Yu et al., 2009),(George et al., 2010),(Wu et al., 2011), (Marcy et al., 2008, Rodriguez et al., 2004).

According to Dee et al (2005), close to ninety percent of physicians employ the use of a PDA device to search for information pertaining to patient related information. Dee et al (2005) also concluded that close to seventy percent of physicians found that PDA is useful and effective to provide relevant information that could assist with the decision making process with patients. It was further reported that using the device also deterred the need to send patients for unnecessary tests and scans, which saves physicians and patients' resources. Furthermore, a minority of physicians also found that patient's diagnosis was changed upon checking PDA device information, and that this had reduced patient's stay in hospitals. These findings were also consistent when compared with trainee physicians, who reported that the PDA had a significant impact in helping to make informed decisions pertaining to their patient's treatment and diagnosis. However, contrary to physicians, more than twenty percent of trainees reported that the PDA helped them to reduce the need for patient's to conduct labs and tests, as well as to change diagnosis for patients (Dee et al., 2005).

Medical library also adapts the use of mobile technology in the library services. Obviously that the capability of the smart phone as a tool to search and retrieve information create a computation between the services providers as well as the medical libraries to design various mobile applications not only for the healthcare professional

but also for the patients as well. number of studies conducted to cover the subject of medical application on the smartphone such as (Payne et al., 2012) ,(Franko and Tirrell, 2011),(Seeger et al., 2015),(Has et al., 2015),(Terry, 2010),(Ozdalga et al., 2012).

2.4 Accessing Medical Information on the Internet.

There are a number of tools have been developed to assist the health professional as well as patients to search and retrieve useful medical information on the internet. Can and Baykal (2007) have divided the medical information access tools into a number of categories shown in figure 2.1.

There are also different medical sources of knowledge including websites and other tools that were used by physicians to retrieve medical information were identified by Leo et al (2006). A total of four thousand five hundred physicians participated in the study. More than ninety percent of physicians reported that they used specific medical databases and sites to retrieve medical information since it is more accurate and relevant to the search query. Although close to fifty percent of physicians also used a regular search engines for general medical information, it was reported that the perception of regular search engines were not as accurate and effective as medical search engines.

Tool	Category	URL
MedSurf	Specialized guide	http://www.medsurf.com
CliniWeb	Specialized guide	http://www.ohsu.edu/clinweb
HardinMD	Specialized guide	http://www.lib.uiowa.edu/hardin/md
MedHelp	Subject directory	http://www.medhelp.org
MedWeb	Subject directory	http://www.medweb.emory.edu/MedWeb
Medical Matrix	Subject directory	http://www.medmatrix.org
MDConsult	Medical resource collection	http://www.medconsult.com
Medscape	Medical resource collection	http://www.medscape.com
PubMed	Medical database search	http://www.pubmedcentral.nih.gov/
MedTexus	Medical database search	http://ai.bpa.arizona.edu/go/medical/MedTexus.html
HelpfulMed	Medical Web search engine	http://ai.bpa.arizona.edu/helpfulmed
MedHunt/Marvin	Medical Web search engine	http://www.hon.ch/Medhunt
WRAPIN	Search on trusted sites	http://www.hon.ch/Project/Hunt_medical.html http://www.wrapin.org/
CMedPort	Medical Web search engine	http://ai.eller.arizona.edu/research/multilingual/cmedport.htm
Mamahealth	Medical web search engine	http://www.mammahealth.com/
Kosmix	Medical web search engine	http://www.kosmix.com
Healthline	Medical web search engine	http://healthline.com
HIQuA	Medical query assistant	http://samwise.bwh.harvard.edu/hiqua/hiqua.html

Figure 2. 1 Medical Information Retrieval tools

There are also different medical sources of knowledge including websites and other tools that were used by physicians to retrieved medical information was identified by Leo et al (2006). A total of four thousand five hundred physicians participated in the study. More than ninety percent of physicians reported that they used specific medical databases and sites to retrieved medical information since it is more accurate and relevant to the search query. Although close to fifty percent of physicians also used a regular search engines for general medical information, it was reported that the perception of regular search engines were not as accurate and effective as medical search engines. The accuracy of medical sites have also been proven through a high number of participants preferring medical websites. Physicians also reportedly used websites and applications such as Uptodate, Medscape, WebMd, Mdconsult and Emedicine (Leo et al 2006).

In addition to the aforementioned methods this chapter will discuss a number of accessing methods that the clinicians used to search for their information, the access methods covered in this section include information retrieval using specialized medical search engines, online medical resources and medical websites, and mobile phones. Along with the barriers they faced while using these accessing methods, whether it was during the patient's consultants or in their daily life (outside the medical organization).

2.4.1 Information Retrieval using Search Engines

Search engines are one of the most important tools on the web, designed to help users locate information related to a particular subject. They involve huge databases, including vast numbers of URL records and HTML. Most information seekers use search engines to retrieve online information, such as Google, AltaVista and many more. Undoubtedly they find both relevant and irrelevant results when they use them to retrieve information (Liu and Gao 2007). Generally speaking, search engines can provide users with the following facilities: (1) gathering all search results requested by users; (2) representing the pages that capture appropriate content; (3) allowing users to send queries; and (4) providing information retrieval methods to refer them to relevant pages (Gordon and Pathak 1999). Nowadays, the majority of medical networks and online medical libraries use search engines to retrieve information and access services (Chen and Dahanayake 2004). In the last decade, health professionals have relied on a number of medical search engines to locate information, such as PubMed, MedSearch, MedlinePlus, WebMD (Lopes and Ribeiro 2011) and Helpfulmed (Chen et al. 2003). The mentioned search engines are specialized in the medical field and designed for health professionals who are experts on performing the search query based on their

medical knowledge and more expert on medical terminologies, these methods may not be helpful for the patients or the users with a limited medical background.

The common goal in using search engine among health professional is for diagnosis and treatment plan. In a recent study conducted by Ajuwon and Popoola (2014) on resident doctors in Nigeria, search engines were proven to be a much used tool for resident doctors. It was found that medical information was primarily accessed via internet resources through medical search engines. Demographic data of the participants were used to determine whether any particular factors would affect the use of online information retrieval. The results of the study found that the demographic factors did not make a difference in the search for information and that the primary search results were pertaining to the need for additional information on a disease/diagnosis/treatment, and also for additional learning (Ajuwon & Popoola 2014). Although that the finding shows that accessing information related to patient care were lower than academic requirements. This shows an evident of the lack of information related to patient care provided by search engines, as well as of the lack of performing the search by the users. The study also reported that there is a need to training on using search engines among doctors to promote the use of internet-based medical information and to increase the accessibility and usage of online sources.

Although the increasing number of specialized search engines in health field nowadays, a study conducted to compare the general and special web search engines in health information retrieval demonstrated that general web search engines surpass the precision of health-specific engines. It also indicated that Google has the best performance, mainly in the top ten results. Wang et al. (2012) reported that the general search engines such as Google, Yahoo!, Bing, and Ask.com are effective search engines

for assisting the user to get health and medical information. This study confirmed of the needs of ranking methods to avoid the large number of the returned results.

As mentioned earlier some of medical search engines may not helpful for users with limited medical experience. Therefore, researchers have designed a number of search engines based medical to assist the medical student as well as the user with no medical knowledge such as MedicPort the system is built based on knowledge of the Unified Medical Language System UMLS (Can and Baykal, 2007). MedicPort provides the user with three main search preferences to perform the search which are: starting with, continuing and ending with. The system has the ability to understand the semantics of web pages and the user queries and based on that the system retrieves the information and present it to the user. When the user submits the query to the MedicPort, the system will transform a keyword search into a conceptual search and searches for related sites and documents related to the search topic and present them to the user to select the appropriate information.

MedSearch, is also a specialized medical search engine that designed by Luo (2009) to help ordinary users who have limited medical terminology and medical background, as well as assisting the clinicians to search for relevant information. This system has a unique feature that allowing the user to extend the length of the queries. It will then reformat them automatically into shorter queries by extracting a subset of important and representative words. This technique of dropping down unimportant words makes the search more significant and also helps to speed up the process and to improve the quality of the search results, especially for the user who are unfamiliar with medical terminology.

MedSearch is another system designed by Luo G et al. (2008) to retrieve the medical information. This system adapts the use of Medical Subject Headings (MeSH) ontology to identify medical terminology and phrase in the returned pages and to order them based on their relevance to the main query. MedSearch interface is shown in figure 2.2.

The results of the medical queries in MedSearch organize into one or more result pages as shown in figure 2.2. Each single page contains ten elements. An element match and link to a web page which contains the title, and the URL of related to the subject. The first page also contains suggested medical phrases listed on the right side of the results page. These medical terminologies are divided into a number of categories (e.g., diseases, treatments, drugs, organs) based on to the classification in the MeSH ontology.

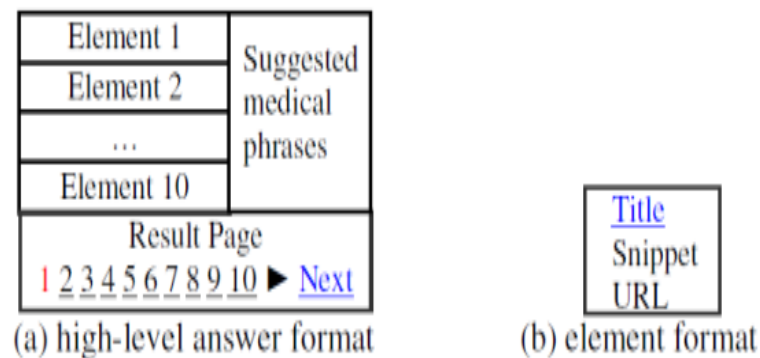


Figure2. 2 Medsaerch Interface (Luo G et al. 2008)

Although, Medsearch designed to help ordinary users who have no medical background, and familiar with using general search engines to search for their information more than other methods. The study has confirmed that MedSearch is

limited on a number of selected medical websites. Undoubtedly, the medical websites that selected for the Medsearch are significant and have a high quality of information. But some medical information queries cannot be answered only through these websites. Another point that is the medical websites may provide some studies and journal articles but not as much as the other medical databases provide, especially some of the databases provide a free access. In addition, the advanced search that other medical libraries provided can assist the user to search for more relevant information rather than posting long queries on the MedSerah.

Another example of the medical search engine is WRAPIN, designed by Gaudinat et al. (2006) to retrieve online health document.

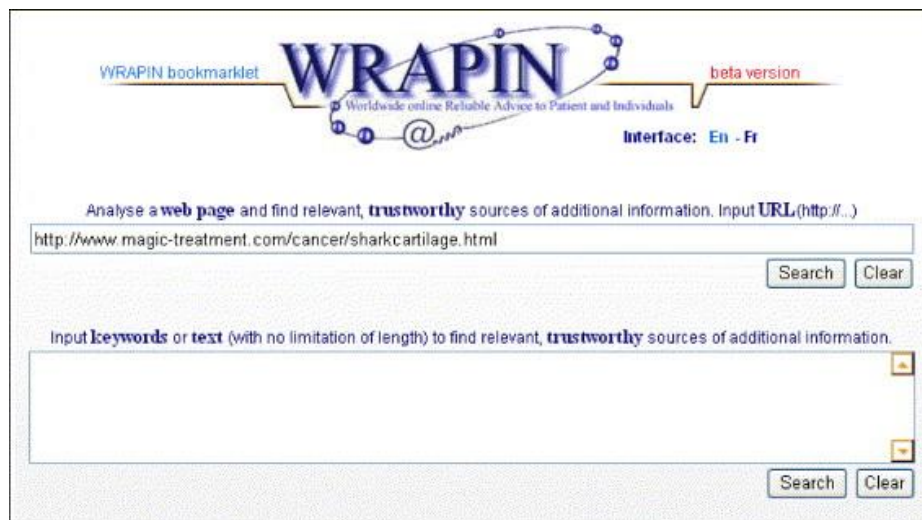


Figure2. 3 WRAPIN Interface

The main goal of WRAPIN was to make the medical information documents available in first place for patients, then for medical professional. The system covers an important range of medical subjects and used a number of medical sources such as PubMed,

Medline, MedHunt and others. Moreover, this system also adapts MeSH terminology to index and retrieve the document. The design of the interface is similar to any simple web search, with the appropriate function that allows the user to search using either URL or normal typing of queries with no limited number of keywords. Figure 2.3 shows the interface of WRAPIN.

The University of Michigan (Hanauer et al., 2015) have developed a search engine called Electronic Medical Record Search Engine (EMERSE) to facilitate information retrieval. The EMERSE has a similar design of google but with a special function such as collaborative search technique that allows the user to manage their queries and save them as search term to be used for the next search or to share them with others. Another feature that EMERSE has, is handling of spelling errors for the medical terminology that contains many difficult to spell the words. Users can also benefit from the query recommendation function, this function provides alternative terms that users can add them to a query these recommendation methods include keywords, synonyms, and a commercial name of the drugs. EMERSE presents the search results via multilevel data views and uses visual cues to assist users to scan the returned documents faster.

In 2009, iMed has been designed by Luo, (2009). It is one of the first intelligent web medical search engines that specifically employs the use of domains and questionnaires to facilitate the need for medical information. iMed designed based on medical knowledge to be used by physicians inside the workplace for the purpose of diagnosis. The system designed a diagnostic decision tree, written by medical professionals. Each diagnostic decision tree is related to either a subjective symptom or objective sign, and each medical phrase under these the symptom or the sign in the diagnostic decision tree

can be used by iMed to generate a query. Figure 2.4 below shows the diagnostic decision tree for a symptom cough on iMed.

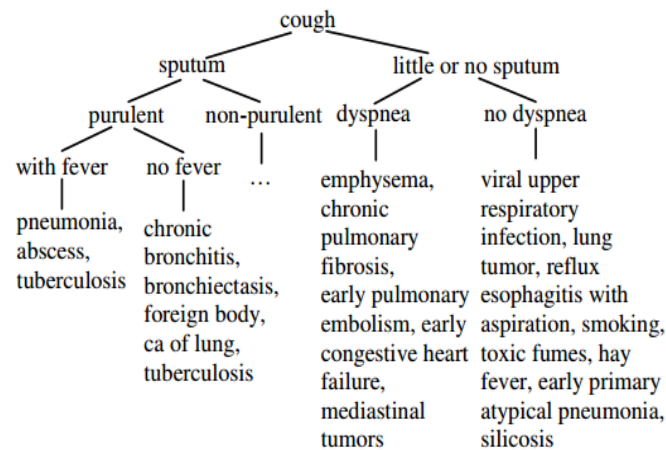


Figure 2.4 Diagnostic Decision Tree

iMed uses diagnostic decision trees to assist the user by generating a number of queries to the user based on the already selected phrase by the user. The user will select one or more symptoms and signs from a list of medical phrase symptoms and signs. Then iMed generates questions related to these selected symptoms and signs. Then, based on the user's answers to the questions, iMed searches the corresponding and then will retrieve a related web pages associated with the disease.

Obviously, iMed is well designed in term of disease categories and the symptoms related to the disease. Moreover, it may help the doctor to retrieve the urgent information during the patient consultant, at the same time it is a time consuming for the doctor to diagnose and search for symptoms and signs.

According to Lopes and Ribeiro (2011), medical search engines are useful in providing relevant medical search results, particularly general health queries, and diagnosis and

treatment options. On the other hand, search engines such as Google have also proven to be effective search methods in producing relevant medical search results, efficiently (Lopes and Ribeiro 2011). The study reported that both of medical search databases and medical search engines should continue to improve the relevancy of search results, since the study concluded that general search engines such as Google was more effective in providing good results. Moreover, medical search engines should also focus on generating results based on the severity of the diagnosis or medical situation, since this can increase the usability of the results (Lopes and Ribeiro 2011).

There are another medical search engine available to retrieve medical information named as Health line which is available on the web for physicians and patients as well (Healthline). Figure 2.5 shows the interface for the Healthline

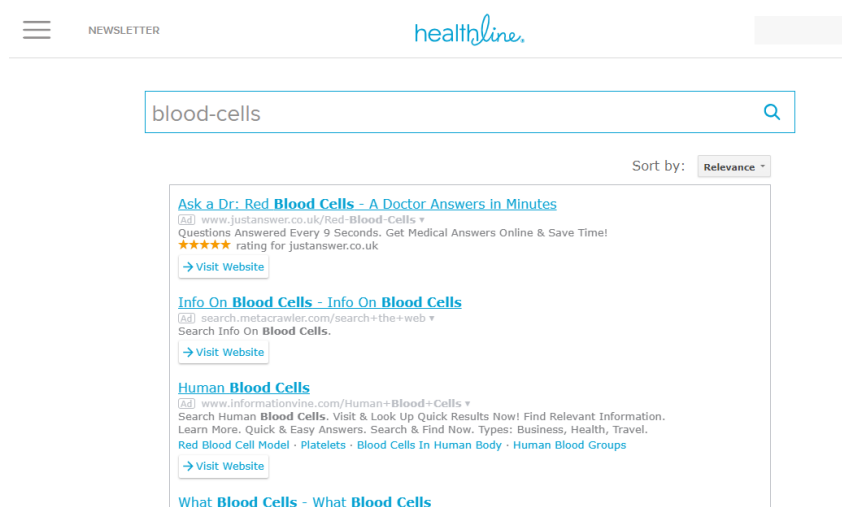


Figure 2.5 Healthline Search Engine

The general search engines such as google can be also used as tool to obtain medical related information. According to Zhang et al. (2012) reported that the general search engines such as Google consider to be a good option for the physicians to search for the information because of the features that the general search offered such as maximizing the number of documents returned for a given request, in addition to that

the general search engines provide more information quicker, compared to medical search engines, although the medical search engines offered more reliable and valid information from trusted sources (Luo et al 2008).

2.4.2 Information Retrieval using Mobile Phone

Smartphones are also another device that has been used in the medical field to access medical information and patient- related data (Ozdalga et al 2012). Smartphones are integrated devices that provide communication options, online data retrieval options as well as allows users to view images pertaining to patient's scans and x-rays, using specific applications designed to improve these images (Ozdalga et al 2012). This saves time for doctors, as they are able to access and provide feedback as well as diagnosis and treatment options for patients using one device. This data can also be transferred to different departments and concerned people, without having to fill up paper works and spend additional time. Smartphones have already become a popular device in the medical industry among physicians and even patients since it is easy to use, easy to carry as well as accessible anywhere. This has shown significant results in terms of productivity and improvement in patient quality care and management (Ozdalga et al 2012).

Although the smartphone has been proven to be effective and significant in improving productivity among physicians in hospitals, there is still a lack of understanding on how this can be achieved using the device (Ozdalga et al 2012). In empirical findings, smartphones have been proven to show an increase in number of patients per physician, and reduced waiting time for patients, however, there are no studies on what applications are used in these phones to gain such positive outcomes

(Ozdalga et al 2012). Furthermore, there is a need to analyse the role of smartphones in the future of the medical industry, since this can change the existing electronic medical systems that store data and records, as well as the impact that it will have towards bridging gaps between physicians and those in the medical field (Ozdalga et al 2012).

Researchers have also begun to focus on the impact of mobile devices and tablets to improve accessibility, and the experience of users as it has better processing systems with a connection to the internet (Chatterley and Chojecki 2010; Franko and Tirrell 2012; Kho et al 2006; Prgomet, Georgiou and Westbrook 2009; Stephens et al 2010; Budiu and Nielsen 2011; Perez et al 2013).

Park and Nam (2009) for instance developed a framework that focuses on integrating medical systems into personalized mobile devices, which could not only enhance user experience, but also function as a usable and relevant medical tool. Besides that, the proposed framework also adapts to the role of the user, providing only information that is usable and relevant to the user rather than the traditional method of providing all information. Moreover, both of the text and image retrieval are improved using wireless networks services and digital imaging software which further improves the function of mobile devices as a tool in assisting medical functions (Park & Nam 2009).

Smartphone operating systems such as iPhone and Android systems have developed a number of applications that allow users to download applications that specifically fits into the operating system, making it highly popular among users including physicians (Franko & Tirrell 2012). The difference with these brands is that it allows users to rate different applications, and provides feedback on the usability of the application. In this case, physicians who find an application useful and applicable in the field, could easily share the information via social media to other people in the field (Franko & Tirrell

2012). Moreover, most applications are free to download, and some are priced very low to increase affordability of the applications (Franko & Tirrell 2012).

This trend is catching up in the medical industry, whereby more app developers are looking into developing more medical specific application that can be used by physicians and nurses. Moreover, there are websites that have been developed to merely cater to review applications so that more people can download and use the application. Although there is a growth in this area, some studies have not focused on which type of applications and operating systems are prevalent in the industry (Abroms et al., 2011), (Dala-Ali et al., 2011),(Franko 2011)(Oehler et al., 2010)(Franko & Tirrell 2012) (Vollmar et al., 2009), (Perez et al., 2013)(Vollmar et al., 2009). It is important for this aspect to be identified in order to create more useful and relevant applications in the industry.

A number of studies conducted regarding accessing libraries using mobile phones, while other studies have addressed mobile information retrieval system in the field of health care. Several researchers have studied accessing library materials using handheld devices, and the use of mobile devices as a tool to deliver library services; these studies showed the most popular services were used among users. At the University of Alberta for example, Carney et al. (2004) found that users who access the library services through the mobile devices and PDA found that both of database search and library catalogue search were the two most popular services among the users, according to the survey question, “Which of the following library materials would you like to be able to download to your mobile/ PDA Respondents were willing to download from the database with 75%, with 46 % for library catalogue. At Washington State University, Cummings et al. (2010) addressed the need for using library services with

small-screen devices such as PDAs and web-enabled cell phones. Their survey found that searching the library OPAC was the most popular service used. Abdul-Karim et al. (2006) investigated the use of mobile phones regarding library and information services at the International Islamic University. Their research found that most respondents were willing to utilise the library services through their mobile phone; they also indicated SMS services which include the reminder of the overdue items, notifications of the new arrival sources, reservations, and query services that can offered by the library through mobile devices.

Jones et al. (2008) report that SMS services can be a good method for the libraries to deliver their services and good tool of accessing learning resources form the libraries Uday Bhaska and Govindarajulu (2008). Mills (2010) also found that users were more positive about accessing information through text messages SMS and were also in favour of being able to access the library catalogue and other services such as “text a librarian” from mobile phone. However, the idea of providing SMS services was not successful at Delta State University, due to the lack of telecommunication infrastructure and networking in the library (Iwhiwhu et al., 2010).

User awareness of using different types of technologies such as mobile phone plays a major role in accessing information. Therefore, users should be aware of how can they use mobile phones to access information especially if they are medical students and the need of information is inevitable for learning and (Iwhiwhu et al., 2010). The low level of awareness and the lack of using mobile phones to access the library services created a kind of competition among libraries to provide a number of services that can be accessed from users’ mobiles, and also to find ways to communicate with their users in an effective manner. Therefore, the communication and an

interaction between the users and the libraries is important if the library wish to remain relevant and improve it services (Popoola, 2008).

There are also several studies have designed an architecture and applications to support multiple-device access including mobile phones, PDAs and iPods. Alvarez-Cavazos et al. (2005) suggested a universal-access architecture for a personal digital library that provides traditional library services such as document submission, full text and metadata indexing and document search and retrieval. The main concern of this architecture is the limitation of the traditional library services. A similar trend made Broussard et al. (2010) through developing a prototype mobile search application for the University of Texas library catalogue. The concern about this effort is similar to the previous study in terms of limitation, but this time in the university community as this application is available for the university staff and students only. Moreover, accessing the library catalogue should not be the only aim for library services. On the other hand, there is some successful framework implementation that should be discussed, made by Castellanos and Sanchez (2003), and aimed at facilitating the generation of interfaces for access to digital library resources from a wide range of mobile devices. Other efforts designed a framework for seamless information retrieval (Han et al. 2006), and an OML framework enabling real-time collection of data on the web (Singh et al. 2005).

2.4.3 Information Retrieval using Online Resources

The use of the online resources (medical databases and medical websites) by the medical professionals has increased in the recent years, a variety of distributed online medical resources on the web such as databases and medical websites are available to many number of users.

Several studies have been conducted on physicians and their usage of electronic sources and information to assist their daily decision making process. The most common source that paid attention among physicians is the online databases such as Medline, PubMed which are primarily used for information retrieval from queries by physicians and nurses as well as patients. There is a growing need for physicians and nurses to use these systems during consultations and when facing patients, due to the increasing complexity of diseases and diagnosis challenges. Furthermore, there is also an increasing awareness on the positive outcome and on the significant results that the system provide for the users. There are however qualms on the accuracy of the system and the slightly difficult aspect of conducting a proper search using these tools (Westbrook et al 2005). According to Westbrook et al (2005), the retrieval of information through online system has proven to be a significant tool that improves the quality of replies posed by participants to a common clinical problem. However, in a small number of cases, the online system may generate wrong answers for the queries and that all depends on the query keywords entered by the user.

There are a number of medical databases available online to the user such as Medline, PubMed, Ovid, Cinal, EMBAS and many more, these databases contain a number of articles from academic journals that cover a number of subjects related to the medicine, nursing, pharmacy, dentistry, veterinary medicine and health care. In addition to the literature in biology and biochemistry (Langton et al 2013). Most of the physicians have rated Medline as one of the more popular and widely used database, followed by PubMed (Leo et al 2006).

McKibbin et al (2007) reported that there are some of the reputable and useful medical resources such as Medline, PubMed, DynaMed. These sources are reliable when it

comes to providing relevant information for physicians to ensure that they are fully understand and able retrieve information related to disease/condition at any time.

As an example of the online databases is **Ovid** it is medical database that has a huge popularity among the Australian's physicians along with Medline and PubMed, this database provides access to online bibliographic databases, academic journals, and other products, chiefly in the area of health sciences. Users can retrieve the information on Ovid based on content analysis using keywords, instead of the primary title. This form of analysis highly depends on the keywords retrieval results that provides matching contents that could be used to further generate medical contents related to recent findings (Elo & Kyngas 2008). The interface of Ovid atabas is presented in figure 2.6.

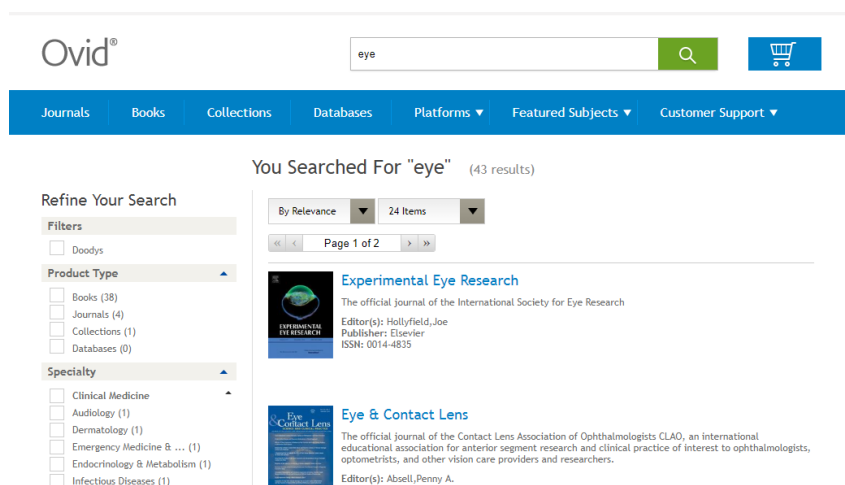


Figure 2.6 Ovid Interface

Another example is **RxNorm** database focuses on pharmacological information for physicians, focusing on specifically identifying and updating drug names according to codes and diseases (RxNorm 2016). The database also acts as an intermediary database that can be used by pharmacies, hospitals and physicians. Each drug is allocated with a unique identifier in order to avoid any confusion among the many similar type of names

and drug components. The primary use of this database is to ensure that all parties gain access to available and new drugs with ease, and at the same time to find newer drug options for diseases.

SnoMed CT is one of the most complex and reliable medical databases that is used by physicians around the world (SnoMed CT 2016). The terms have been developed by the medical organization that focuses on developing medical terminologies using international standards. The purpose of this is to standardized key terminologies, which makes it easier to search and find information via medical databases. The database is also integrated with electronic health record (EHR) systems in hospitals, which provides better accessibility for physicians who are in need of searching for medical information. This also allows the system to be used during consultation in order to provide immediate explanation and options. The database also provides multi languages, which removes language barriers.

In addition to the aforementioned databases there are also a number of medical databases available online to the users. Table 2.1 covered some of the common medical databases available for physicians, researchers, students, and patients.

Medical libraries websites are another type of online resources that designed to assist physicians, health professionals, students, patients, and information specialists in finding health and scientific information to improve, update, assess, or evaluate health care. This source of medical information is a popular source that used by physicians to seek for information pertaining to diagnosis and treatments via various devices such as Pc, mobile phone and personal digital assistant (PDA).

Table 2.1 Medical Databases

Database	Description
Medline	It is a huge bibliographic database of life sciences and biomedical information. It includes bibliographic information for articles and literature for medical sciences
PubMed	It is a database of reference and abstracts on life sciences and biomedical topics.
CINAHL	Database covers a wide range of topics including nursing, biomedicine, health sciences librarianship, alternative/complementary medicine, consumer health and 17 allied health disciplines
ASSIA	Covers health, social services, psychology, sociology, economics, politics, race relations, and education.
EMBASE	Database covers all medical drugs and disease from 1974 to present.
DOS	Dentistry & Oral Science Source is a full-text research database for dental practitioners and researchers.
IPH	International Pharmaceutical Abstracts is a comprehensive database provides indexing and abstracts for pharmaceutical and medical journals.
TOXNET	Is a medical database covers toxicology, hazardous chemicals, and biohazards.
AMED	Allied and Complementary Medicine Database.
Global Health	It is a database for public health literature

The primary function of medical websites are to optimize on information organization and provide precise search results that specifically caters to the search request, as quick as possible. However, some of the search results may result in redundancy which causes waste of time for physicians as they have to go through page by page to access relevant and required information (Luo et al 2008).

There are a number of medical libraries available online such as the Medical Library Association (MLA), National Institutes of Health, National Library of Medicine, British Medical Association (BMA) American Medical Informatics Association (AMIA), Association of Academic Health Sciences Libraries(AAHSL) all considered to be a

good examples of online library that present a biomedical and health information to the user.

The last online source that used by the physicians to share information is the social media. Nowadays physicians are able to use the social media as an application and platform to share useful medical information with other physicians, patients, as well as the general public. According to McGowan et al (2012) in their study which focused on the usefulness of the social media technology, and the attitude of physicians towards embracing it, a total four hundred and eighty five physicians participated in the study, only twenty four percent of physicians used social media on a daily basis, close to fifteen percent of physicians contributed medical related information or personal information via social media applications on a daily basis and the majority of them participated in social media on a weekly basis, and less than half of the participants participated in social media.

Physicians had a positive perception of social media, stating that it is a useful tool to interact with others, and to get good information. Physicians also stated that social media allowed them and patients relationships to develop, which also increased the quality of patients care and management (McGowan et al 2012). However, social media is a useful and effective tool that can be used to achieve many positive outcomes, in contrast, social media in general has a negative image when it comes to providing and retrieving trustworthy data due to the nature of the platform that allows everyone to give their opinion (McGowan et al 2012).

2.5 Factors and Barriers effect seeking medical information.

- **Gender and location**

There are two types of factors that have an impact on how physicians search for information which are internal and external factors. The internal or the personal factors such as age, gender and specialization does have an impact towards how and when physicians search for information (Gonzalez-Gonzalez et al 2007). Okoro (2009), found that more than ninety percent of male physicians retrieved and searched information primarily for self-development purposes, and to update latest findings in their specialization. On the other hand, female physicians retrieved information to prepare for exams, to update their knowledge and to focus on long term career development.

On the other hand, external factors such as location of practice also impacts how physicians can search for information. A study conducted by Andrew et al (2005) and Tumwikirize et al (2008) found that physicians who had practices in rural areas were less likely to get access to online information, and hence relied on books and existing information through experiences whereas physicians who were located in cities and urban areas were more likely to access information via medical websites and journals.

Davis (2011), confirmed that the location of practice plays a part in how physicians search for the information, by comparing a different methods that physicians in the United States, United Kingdom and Canada used to enrich their medical knowledge and retrieved information. The study found that most British and American physicians were more likely to rely on online databases compared to Canadian physicians.

- **specialization**

Other factors such as specialization will also change the need for physicians to seek for information. For instance, general practitioners are more likely to seek for updated information regularly compared to a paediatrician (Bryant 2004; McConaghy 2006).

- **Devices**

Other factor is that the type of devices that used to access the information. Personal digital assistant (PDA) has been an essential part for health professionals and physicians to search for information pertaining to patient and other medical queries (Fox et al. 2007). However, applications such as Ovid, MEDLINE and PubMed were unpopular searches on PDAs, compared to PEPID and up to-date as the electronic versions made it easier for users to search for information required (Chatterley & Chojecki 2010). Hence, it is important for the applications to be easy to use and to provide relevant information pertaining to the search query in the fastest time as possible. Furthermore, physicians also found that PDAs helped to increase efficiency as well as helped to provide better decision making options, which also had a positive correlation with the management and consultation of patients (McCord et al. 2009). On the other hand, there are also arguments that physicians who spend time typing on a device may actually have the opposite effect towards patients who are expecting a more engaging conversation as the usage of devices, may reduce face time with physicians (Margalit et al. 2006).

McCord et al (2009) also found that physicians reported positive improvements in the ability to access various information pertaining to medications, and diagnosis related queries through their mobile devices and PC. Hence, the usage of a device may have a

positive or a negative impact towards patients, and this varies with the patients and the physicians.

- **Patient and physician interaction**

The quality of patient and physician interaction was reported to be an important factor that affects seeking information. Patients often prefer to have a personal conversation with doctors, without the use of any gadgets or devices (Mollon et al 2009; Moxey et al 2010; Pearson et al 2009). Physicians on the other hand, need to access some information in order to get in to the final decision for the patients. Other studies also have confirmed that the interactions between patients and physicians have a positive therapeutic benefits in the patients outcome process (Boonstra and Broekhuis 2010., McCord et al. 2009). In many cases, patients are likely to be more satisfied with their physicians if they were able to have an engaging conversation and if the doctor were empathetic towards their diagnosis. In order for this to happen, physicians need to have sufficient information pertaining to the patient's diagnosis and treatment plans.

- **Accessibility**

Having an accessible source of information is important to ensure that there are no barriers that prevent the interaction with, or access to the sources. The online medical databases such as Medline have gained popularity among physicians, primarily due to the fact that it is accessible from any location and any time, whether was during consultation time with patients or while accessing recent information that can significantly improve physician and patient outcome (Brandt 2012; Kohane 2012; Mollon et al 2009; Moxey et al 2010; Pearson et al 2009). However, there are some of the challenges identified are pertaining to the payment and subscription and some

related to the difficulty in accessing and searching for information, as this varies with different medical service providers (Moxey et al 2010).

- **Lack of time and awareness**

Evidence based medicine is an approach that focuses on finding the most appropriate evidence before making a medical decision in order to determine the existing possible options that are available for patients (Sackett 2005). However, this is not the case in most medical organizations due to time constraints and lack of training and awareness on this issue. According to Nwagwu and Oshiname (2009), many doctors are not aware of the existing medical websites that could help to provide options and explanations for diseases. Moreover, doctors often have a lack time since there are high number of patients during consultation, which makes it difficult to search information for individual patients (Barghout et al 2009). Besides that, Vliveness et al (2009) reported that doctors often lack skills in searching for information, and should be trained to properly utilize existing resources.

Physicians also face challenges finding relevant resources that are properly organized which can save valuable time, especially during consultation with patient's (Agency for Healthcare Research and Quality 2014).

Furthermore, time consuming is important while searching the information especially in an environment like critical care units, information required are often intensive and thorough, that needs to be accessible instantly in order to be useful in the situation (Patel et al 2008., (Amsbary & Powell 2013), (Bates et al.2006., Falagas et al. 2008).

2.6 Tools for Indexing Medical Information

In the medical field there are two main important sources (MeSh and UMLS) that used to index the medical terminology and both are widely used in most of the medical retrieval systems such as Medline and PubMed this section will cover them.

2.6.1 MeSH

MeSH (Medical Subject Headings) is a database that contains a huge number of medical terminology and thesaurus, organized alphabetically and using hierarchical structures (MeSH 2016). The hierarchical structure focuses on broad terminologies while lower level hierarchy findings comprises of more specific illnesses and terminologies.

The MeSH structure starts from a generic headings such as “Body Regions” and goes down to more specific heading and it goes then deep in the structure that specify a specific subheading headings such as “Ear” or “Face”. Each heading in MeSH is represented by number called tree number. For example, the heading “Head” is represented by the MeSH tree number “A01.456”, and its sub-heading, “Ear” “A01.456.313” and “Face”, “A01.456.313” as shown in figure 2.7 (Alghoson, 2014)

The database is suggested by the United States National Library of Medicine (NLM) to index medical articles that are often published in PubMed and Medline databases as well as other medical databases (MeSH 2016).

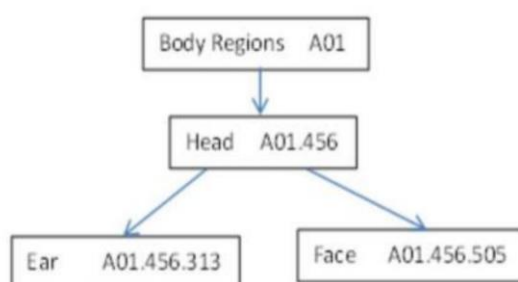


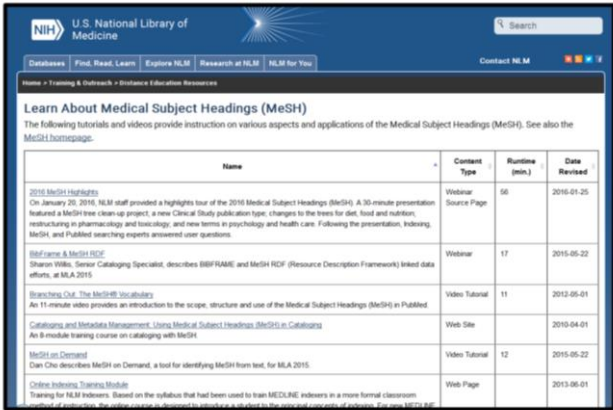
Figure 2.7 MeSH Structure

MeSH vocabulary reflects the progress in medical sciences and the growth of the medical vocabularies, these new vocabulary are added to the MeSH terms and the already existing terminologies are modified and updated each year (MeSH 2016). MeSH is used by several health organizations as well as medical libraries, including MEDLINE, to organize materials and index information. Moreover, most of the medical search engines and retrieval systems that mentioned in the literature are considered MeSH as a main source of indexing and organizing the medical subjects.

MeSH data can be accessed via the website and can also be downloaded in different formats for convenience. The database also contains fact sheets and basic guidelines and tutorials on how to search using the database, with a step by step guideline to assist users. Apart from that, the website also provides links to medical publications (MeSH 2016).

Figure 2.8 below shows a screen shot of information pertaining to the MeSH database, where there are different content types provided such as webinar, videos, websites as well as tutorials in different formats. The format of the website is quite

user-friendly, with most of the instructions and guidelines provided in the first page. The efficiency of the system is shown as the time of the content provided helps to give physicians and users a time frame of what to expect when they click on a content. The dates on the other hand, act as a guideline as to when the content was published or produced.



Name	Content Type	Runtime (min.)	Date Revised
2016 MeSH Highlights On January 20, 2016, NLM staff provided a highlights tour of the 2016 Medical Subject Headings (MeSH). A 30-minute presentation featured a MeSH tree clean up project, a new Clinical Study publication type, changes to the trees for diet, food and nutrition, restructuring in pharmacology and toxicology, and new terms in psychology and health care. Following the presentation, indexing, MeSH, and PubMed searching experts answered user questions.	Webinar Source Page	56	2016-01-25
MeSHtree & MeSH RDF Sharon Wilks, Senior Cataloging Specialist, describes BIBFRAME and MeSH RDF (Resource Description Framework) linked data efforts, at MLA 2015.	Webinar	17	2015-05-22
Branching Out: The MeSH-99 Vocabulary An 11-minute video provides an introduction to the scope, structure and use of the Medical Subject Headings (MeSH) in PubMed.	Video Tutorial	11	2012-06-01
Cataloging and Metadata Management: Using Medical Subject Headings (MeSH) in Catalogs An 8-minute training course on cataloging with MeSH.	Web Site		2010-04-01
MeSH on Demand Dan Cho describes MeSH on Demand, a tool for identifying MeSH from text, for MLA 2015.	Video Tutorial	12	2015-05-22
Online Indexing Training Module Training for NLM indexers. Based on the syllabus that had been used to train MEDLINE indexers in a more formal classroom. Includes a glossary, the online version of the MeSH tree, a list of MeSH terms, and a list of MeSH terms that are not in the MeSH tree.	Web Page		2013-06-01

Figure 2.8 MeSH Database (MeSH 2016)

Figure 2.9 on the other hand, shows a glimpse of the training materials and guidelines provided to assist users to maximize the search database (MeSH 2016). As shown below, the guidelines are divided into eight different modules, which allow the users to break up the training period if required, and this can be continued when required. Furthermore, the organization of information is clear and concise, which allows users to understand and follow the process efficiently. However on the down side, there are too many topics to be covered in order to fully understand and optimize the database due to its large availability of medical data and guidelines. This may definitely cause challenges for physicians to access, since it requires time to explore and to fully get used to the system.

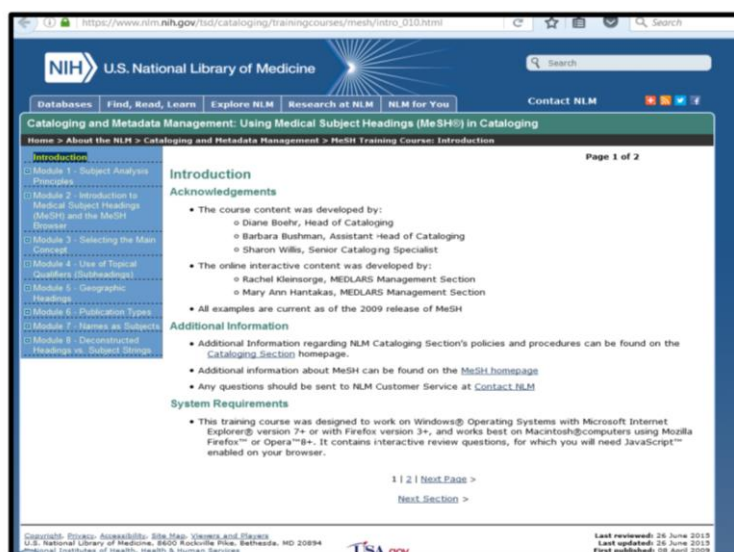


Figure 2.9 MeSH Training (MeSH 2016)

2.6.2 UMLS (Unified Medical Language System)

The Unified Medical Language System (UMLS) is a database that integrates and uses key terms and applies the international coding standards to create a more efficient database of electronic health records (Unified Medical Language System, 2016). The database also promotes a better interoperable system that is highly effective and efficient in distributing medical information. This database was created to assist the integration and transfer of information to computer systems, which could simplify the process of making computers understand medical terminologies (Unified Medical Language System, 2016). Furthermore, specific software and tools are used by the National Library of Medicine to develop and to improve the process of finding and integrating biomedical terminologies and health related information online. This database is also multi-purpose as it caters to developers and medical users. Developers are able to access information of patient records, guidelines as well as health data, while maintaining the ability to modify and customize the system according to new needs

(Unified Medical Language System, 2016). Apart from that, the UMLS services allow users to access the site through web interfaces and web services to access data (Unified Medical Language System, 2016). Some of the file formats are from UMLS knowledge sources, RxNorm especially for receiving weekly or monthly updates, SNOMED CT and administrative data (Unified Medical Language System, 2016).

There are three primary tools and sources in the UMLS which are the metathesaurus, the semantic network and the specialist lexicon which is updated twice a year to ensure up to date information distribution (Unified Medical Language System, 2016). The metathesaurus is a database that contains medical vocabulary in different languages that can be changed according to the need and user, and it can also be stored and retrieved using the meaning of the word. Some of the terms included are from the areas of health services, health statistics and patients care (Unified Medical Language System, 2016). There is more than a hundred and fifty vocabulary systems such as SNOMED CT and MeSH (Unified Medical Language System, 2016). Two primary formats are used which is the Rich Release Format (RRF) and the Original Release Format (ORF).

The semantic network on the other hand, focuses on broad medical categories and useful information between two semantic relations such as diagnosis and treatment key terms (Unified Medical Language System, 2016). Besides that, the specialist lexicon focuses on providing users with the information required for the natural language processing system (Unified Medical Language System, 2016). This system is generally run in the English language, and consists of an English medical dictionary with medical and biomedical terminologies. Furthermore, this tool is run using the Java program which is intended to assist users to better manage the variation of search and

results in the biomedical context. The generated terminologies and data are as a result of information retrieved from databases to index NLP applications (Unified Medical Language System, 2016). Users are able to access the UMLS within the United States for free, and this also applies to international users, which makes it a widely used medical tool (Unified Medical Language System, 2016). However, the usage of the system requires a quick sign up especially for the usage of metathesaurus due to some words being protected by some producers. Specific content viewing requires special permission from certain producers, often free of charge (Unified Medical Language System, 2016).

2.6.3 ICD-10 Codes for Information Retrieval

The classification of diseases is also known as "nosology", which focuses on keeping track on different diseases and identifying the causes of death for people (O'Malley et al 2005). Nosology became a primary concern when there was a rising need for medical assistance through insurance companies, which demanded the need to check for medical details such as diagnosis and cause of death (O'Malley et al 2005). Some of the existing code of classification for diseases includes the International Classification of Diseases (ICD), the American Medical Association Current Procedural Terminology (CPT 4), Medicaid and Medicare services, the American Psychiatric Association's Diagnostic and Statistical Manual of Mental Disorders (DSM-IV) and the Agency for Healthcare Research and Quality's Clinical Classification Software (CCS). The International Classification of Diseases (ICD) is a code that is used by those in the medical industry, particularly physicians and nurses as a diagnostic tool (WHO 2015).

For medical researchers such as physicians, the ICD codes are important as it provides structured information that is useful when conducting research in the area of medicine. Patients are often grouped according to their diagnosis, which require the use of ICD codes in order to identify the patterns of the disease, the care strategies and methods for dealing with the disease as well as the possible outcomes of the disease (O'Malley et al 2005). Furthermore, ICD codes also provide medical researchers to analyse risks, and to compare cross-sectional findings and studies in order to explore different aspects of the disease. The primary motive is to ensure that the patient gains the best possible care strategy. Besides that, researchers often refer to the ICD codes to create inclusion and exclusion criteria for the proposed population of samples in their studies. These criteria help to set a certain boundary in the study which makes it a good quality study. These codes also allow for studies to document the complications faced by patients, to track the rate of utilization and process, as well as to determine the reason behind the morbidity and fatality rates for diseases. Hence, the ICD codes can be classified as an important source of the medical field due to its diverse applicability and use in the industry (O'Malley et al 2005).

The classification contains many diseases, diagnosis and terms that is often used in studies by researchers, as well as to search in medical databases. These codes are often expanded to include more causes of death for purposes of health researchers, reimbursement and for administrative purposes (O'Malley et al 2005). Furthermore, the ICD database also contains a lot of information pertaining to different types of diseases that are in populations in an organized manner, using codes to create a database whereby the information can be easily accessed from anywhere (WHO 2015). Furthermore, ICD databases are often also used for training purposes. Some of the sections include codes, death certifications, ethics issues and privacy considerations,

quality issues, statistics, as well as general knowledge (WHO 2015). There are up to 10 versions of the ICD codes, which are often updated when it is required, or when there is a need to fill up a gap in the medical literature (WHO 2015). Each version may contain slight variations, which could be checked on the World Health Organization website for clarification. The codes are also updated, and new codes are created to keep up with the expanding and new diseases and diagnosis. Therefore, this ensures that a person in the medical profession is always kept up to date with recent additions. Figure 2. 10 is a screenshot from the World Health Organization website, which shows the latest ICD 10 version that was updated in 2016 (WHO, 2016). The left hand side of the website shows some of the different conditions and diseases, whereby a user can click on to find out additional details pertaining to the query searched.

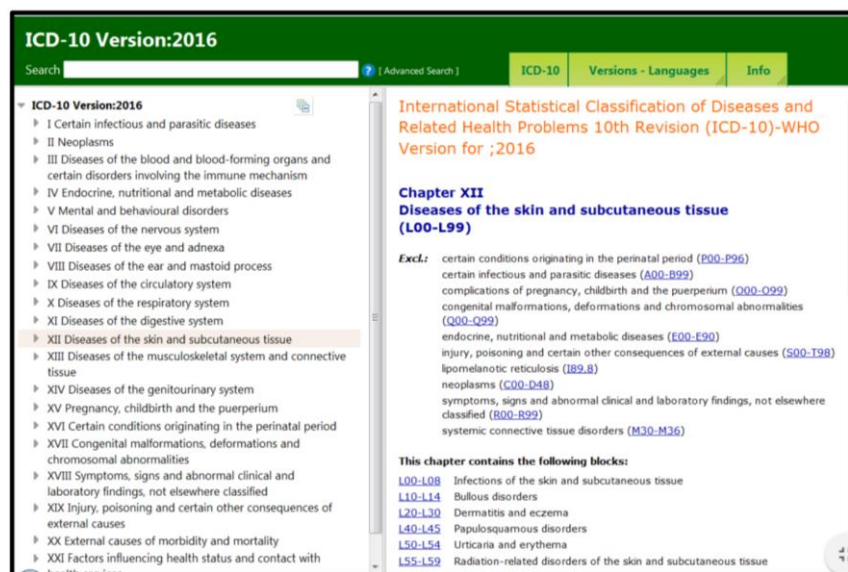


Figure 2.10 ICD 10 (WHO 2015)

On the right hand side, the different chapters are expanded to include definitions, medical information as well as a breakdown of sub-headings that comes under that particular category. This shows a very organized and structured manner of

accessing information, which also makes it easier for physicians around the world to access and understand (WHO 2015). Furthermore, the website also includes a brief training manual on how to search for specific information, catering to those who are still in training and who would like to have access to these information. This database can also be used as a training tool for young physicians and nurses who are getting used to medical databases, since it is easier to find for information. Apart from that, the ICD 10 also includes different languages, for the benefit of physicians who do not use English as the first language (WHO 2015).

A study was conducted by O'Malley et al (2005) to identify the probability of errors in the ICD coding process. The study used a systematic review method to summarize some of the diagnostic codes that are used from the point of admission of patients, to diagnosing the disease through codes. This is an important area to analyse since the ICD often used to classify death and death related information, and hence this is an important area that should not have errors, or should have limited margin for error. The study concluded that there were errors primarily in the area of patient care and admission, which primarily consisted of information such as point of admission, the different communication occurrence between physicians and patients as well as the physicians' knowledge and prior experience with the diagnosed illness. Coders should also pay more attention to smaller details in order to be aware of potential mistakes and errors, as well as to evaluate the applicability of the codes generated (O'Malley et al 2005).

2.7 Important Issues outlined the proposed solution

- 1- Some of the existing solutions have discussed ranking results but a few of them have addressed user needs and ranking results based on the user profile or user preference. Therefore, there is a need to identify the user needs and draw requirement from the existing methods. This is the prime motivation for the re-search work. We intend to propose a generic framework that has ability to personalize the search results based on the user profile.
- 2- The literature has classified the medical information retrieval methods into 3 main types medical information retrieval using databases, web based medical information and mobile based application. This classification helped to design the generic framework to adapt different type devices such as PC, Smart phone, and PDA.
- 3- The user of medical information retrieval system are divided into two main types general and different user from medical domain this division helped to identify the expert user this will help to design the smart layer that will address the results customization.
- 4- Finally, to provide an operability, to the generic framework, there is a need to design a protocol that will performed signalling and data transfer for the user' request.

Despite attempts by researchers to provide medical information retrieval systems that discussed in this chapter, these efforts need to be further developed. In spite of the success of some attempts, the number of problems identified correlates to the amount of disorganised information and the huge number of distributed databases. The main problem identified is that some search systems are very generic in terms of information

retrieval and resemble general search engines, such as Google. Another concern is that although some retrieval systems have a specific engine for medical searches, other studies have demonstrated that general search engines are limited and do not always provide accurate information. Attempts made by researchers who used mobile phones as tools for accessing information also showed a lack of structure and a limitation on specific services, such as SMS and library catalogue search service. Therefore, there is a need to design a universal framework to allow the user to access to the medical information and allow new services and applications related to medical libraries systems to be introduced.

The primer focus on the research is to design a framework that solves the core problems mentioned above and to address the problem of customising and personalising the search results.

2.8 Summary

This literature review chapter focused on exploring existing literature on medical libraries and databases, as well as the retrieval systems in place. A brief introduction of the health care system and the need for information was discussed (Luo et al 2008; Johnson 2008; Gonzalez-Gonzalez et al. 2007; Altschuler et al. 2012). Some of the barriers and challenges faced in the healthcare systems were also discussed, particularly focusing on the issue of lack of information, and difficulty in accessing relevant information. This was followed by a discussion on technological advancements in relation to the information retrieval systems, and its impact towards the healthcare system. Several studies were also explored, in relation to the different devices and retrieval methods that were reported. The second part of the study focused on the need

for medical information sources, which focused on the information seeking behaviour among physicians and the need to identify sources through the usage of different devices such as PDA and social media to share and exchange information. Next, the literature focused on how physicians searched for medical information. This explored the many ways that are used by physicians to find information, particularly during consultation with patients and for diagnosis and treatment purposes. In the last part, the focus was on the search tools that were used, from TREC systems to Medline systems. The importance of the ICD codes were also discussed in relation to retrieve relevant information. Lastly, the focus was on the step by step process of searching for information on the medical database.

Chapter 3: Methodology and the

Proposed Framework

In this chapter the framework designed methodology will be presented. This chapter presents an overview of the concept and the design methodology of the framework and the objectives that the framework aims to achieve. Thereafter, the chapter will introduce to the reader the layered architecture of the framework that related to the scope of this study and its main components, together with a description of their functionality, will also be presented. It will then give a detailed description of protocols and interfaces used to link each units in the framework. Moreover, it will present the user profile along with the user identity and finally will illustrate how to request and update the user profile.

3.1 Research Methodology

The methodology of the research followed in this research is based on the methodology found in the literature such as the health monitoring system by Rikitake et al(2009) ,Mobile Telemedicine System by Navarro et al (2006), mobile-Health Care System by Mayuri Gund et al (2011), and Medical Sensor Application Framework by Markota & Ivica (2010).

The logical methodology flow since the beginning of carried research can be summarized in the diagram shows in figure 3.1 each step in the methodology followed will be explained in the following subsection.

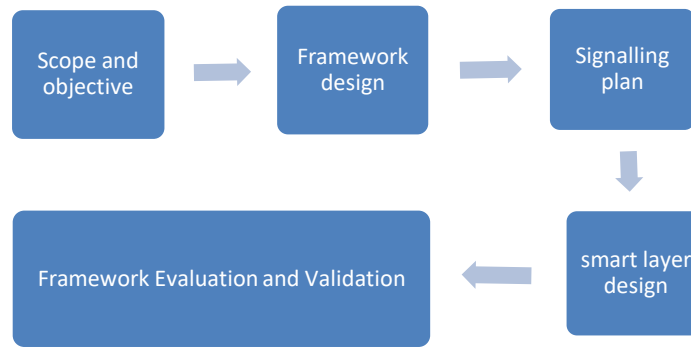


Figure 3.1 Research Methodology General Flow

3.2 Scope and Objective Methodology

The designed framework takes advantage of the IMS architecture and adapts the layered model. The new framework (MICF) designed to provide seamless connectivity for a different type of users and different methods of access for that reason the access layer is designed, and for the purpose of the connectivity between the user and the Application Service (AS), the control layer is created positioned logically above the access layer. The control layer controls the signalling in the framework. The application layer located above the control layer. The separation between the access layer and the application layer is essential to allow the users messages to go through a number of a process before reaching to the application server, and that is an essential part in communication services for service integrity it is also important to simplify the design and operation of each layer via having a set of interfaces and protocols for the signalling plan that ensures smooth transmission between different units in the framework.

In this study the main focus will be on the main connectivity layer and that start by given the reader a detailed description about the main components and the function for each unit, the signalling plan for the message which controlled by number of protocol and interfaces, and will also focus on the smart layer which design for the purpose of customizing the user search. However, this study will not focus on the AS and because it deals with high-level service management in which advance services are delivered and the application services are provided by the services providers without any impact on the access and the control layer which keeps establishing monitoring the signalling without any harms.

The transaction in the system will be performed using SIP protocol. Which plays a major role in the system for initiating, maintaining, modifying and terminating the sessions, and because MICF system designed for medical professionals any delay for the responses at any transactions done by the physicians may affect the patient's life. The delay of the response time also has a negative impact of the Quality of Services (QoS). Therefore, it is important to evaluate the system performance in order to give an indication of the overall system capacity and scalability which will be affected when a number of users accessing the system at the same time to request information. And that can happen during the consulting time with a huge number of patients or can be happened at the emergency unit when the needs of urgent information for design and treatment plan is critical for their life. Therefore the evaluation will focus only on the response time of the request which reflects the scalability and reliability of the overall system.

The methodology for the research can be summarised as follows:

- Determine the current medical information access methods and determine the pros and cons of each method.
- Determine the needs and the challenges that led to designing the framework (MICF).
- Design a comprehensive framework and present the architectural layers and define the main functions of each layer.

- Define the framework functionalities including the connectivity (protocols and interfaces) and signalling plan for the message
- Identify the needs for the smart layer and designed the smart layer
- Determine the appropriate tool for the evaluation to generate the results
- Determine the evaluation performance metrics to validate the system performance
- Generate the results and write the evaluation conclusion.

3.3 Framework Design Methodology

The proposed framework designed is based on the current deployed communication system IMS system performance in the market. According to the literature, there are challenges associated to the current access methods when it comes to providing scalable and reliable system for different users and different access devices.

The methodology followed during the framework design stages take into account the user profile and customization as consequent, the smart layer is designed and presented as a part of the control layer.

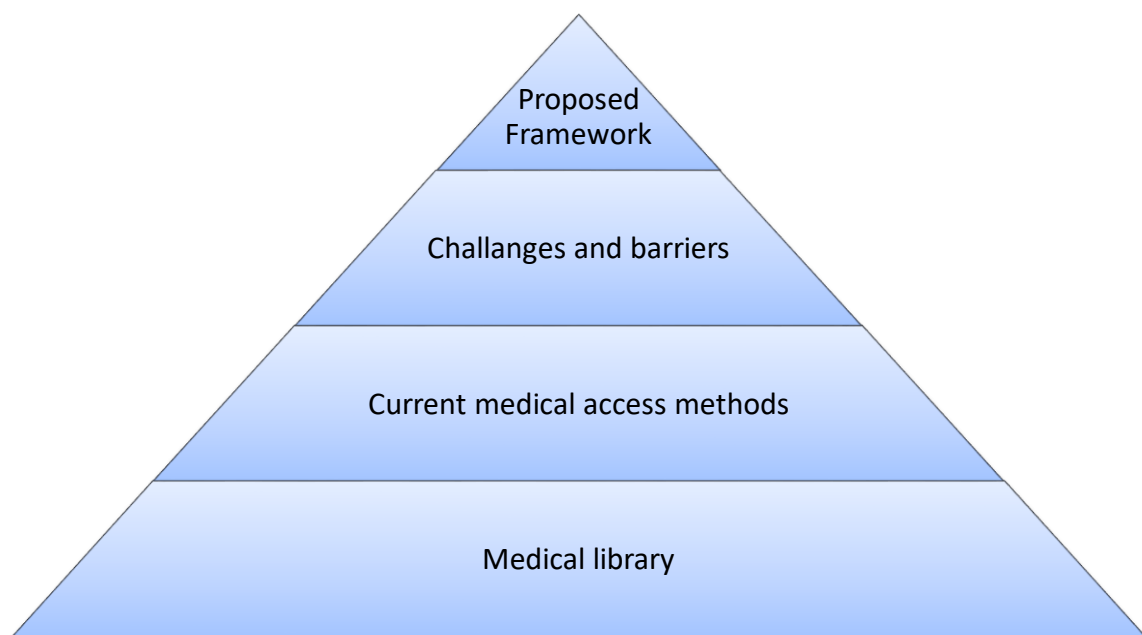


Figure 3.2 Organizational Diagram of the Framework Design

Figure 3.2 shows an organizational diagram of framework design methodology structure.

The structure represents the logical flow of the methodology to design the framework. The vision of medical libraries tend to be in general very optimistic as showed in a number of studies in the literature as the trend toward designing medical information system such as (Ovid, RxNorm, SnoMed) and medical search engines such as (MedSearch, iMed, MedicPort, WRAPIN). Studies summarized in the literature helped in identifying the general research trend and efforts paid toward filling the identified gap. The efforts made on accessing the medical library through a mobile device create a kind of challenge for this study since that mobile devices allow an easy access to the information at any time/place. Through a deep studying of that technology, this challenge has shaped the way of designing the framework by adapting the current communication model IMS within the allocated time frame.

It worth mention that the IMS is a multimedia communication system, is an open system architecture that supports a range of IP-based services over network. It enables peer-to-peer real time services, such as voice and video. It has a common session control layer based Session Initiation Protocol (SIP) which plays the main role for setting up and controlling the session. This system has been adapted to design the layer architecture.

It is important to note that the IMS is not really customized to work for medical' users, it is mainly designed for multimedia application in commercial mobile networking. . However, there is a big different between the IMS system and the proposed framework

- The IMS is a communication system that supports voice video and radio communication. It is designed to provide robust multimedia and live streaming services across roaming boundaries and over diverse access technologies such as (wireless, broadband, and the fixed line)

- The IMS is a huge system that supports different access technologies such as 3G, 4G and 5G.
- The IMS has more than one gateway, it contains three gateways (MGCF, SGW, and MGW) an integrity session control function to identify the network dominoes for different users and more than one server and databases.

The new framework (MICF) is a retrieval system designed to retrieve medical information. It adapts the IMS layered model, it inspired by the IMS and customized to be more adaptable for the needs of the medical staff by adding the smart layer which is mainly catered to the purpose of customizing the user 'search results and that is a new contribution compared to another system

- The system required an internet access.
- The MICF framework uses the proxy as gateway to the system because there is only one domain which is the medical domain.
- There is only one server and one databases for the user.
- The smart layer add unique value to the new framework and because of the special function that the SL performs it has to be located inside the control layer.

3.4 Introduction to the Framework

By considering the previous studies mentioned in the second part of this report, this research identified the weaknesses and strengths of previous attempts. This research proposes a new system as a possible solution to the earlier identified problems listed in chapter 2. The researcher also takes into account the advantages of some attempts that were mentioned before.

This research aims to design a generic framework for accessing medical libraries, offering the flexibility to achieve the following design objectives:

- Allows multiple methods of accessing information, such as mobile, PC, PDA, diagnostic tool and other devices.
- Offers a scalable system for a large number of subscribers.
- Has the capability of handling adaptable content and methods of access.
- Allows new applications and services to be introduced with no requirement for major changes.
- Allows access to information via different type of applications.
- Allows multimedia services (video, voice, audio) to be presented.

The new framework is designed based on a layered architecture similar to that used in the communications domain for the design of the IP Multimedia Subsystem (IMS). IMS is standardized by the telecommunications world based on new concepts that provides a real-time multimedia sessions (voice session, video session, conference session, etc.) and non-real-time multimedia sessions (Push to talk, Presence, instant messaging) over an all-IP network. IMS targets convergence of services supplied in differently by different types of networks: fixed, mobile, Internet (Simon and jean 2013).

The IMS layered architecture is shown in figure 3.3.

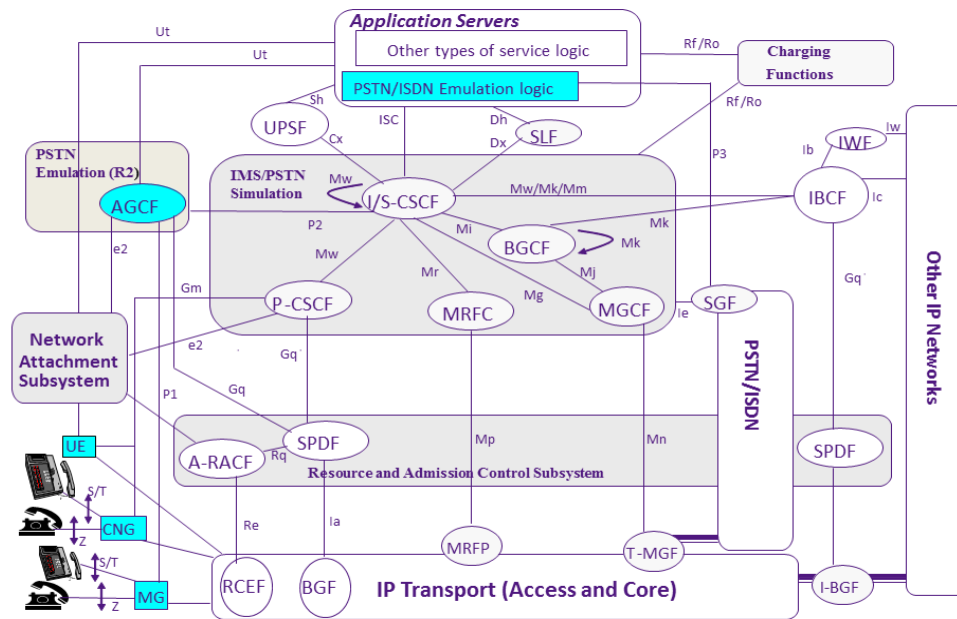


Figure 3.3 IMS Architecture

The proposed architecture called MICF shown in figure 3.4 consists of four main horizontal layers: access/user layer, control layer, smart layer and application layer. The user layer consists of different types of users (medical staff from different specialties as well as patients) and different entry devices. The control layer is a SIP server that hosts the core system Medical Information Control Function (MICF). It incorporates the application layer, and the entry users' devices and it is responsible for connectivity and controlling entry. Moreover, it has three elements: the Proxy Medical Information Control Function (P-MICF) and the Serving Medical Information Control Function (S-MICF), which are responsible for authorisation, identification and routing requests to the application services, and the Medical Subscriber Database (MSDB), which holds user profiles.

The third layer is the intelligent layer; this layer has been created to add value to the framework in terms of customising and filtering search results, and is considered to be the main feature of the framework. Moreover, this layer provides the user with an alternative search method, which will be described later in this chapter. Finally, the service layer

comprises application and content servers and executes value-added services for the user. This layer consists of different types of services provided by the medical libraries, such as databases, e-journals and other services.

3.5 Framework Architecture

This section gives a brief overview of the layered architecture and a detailed description of the main components and their functions, in addition to a description of the manner in which way that the framework will function.

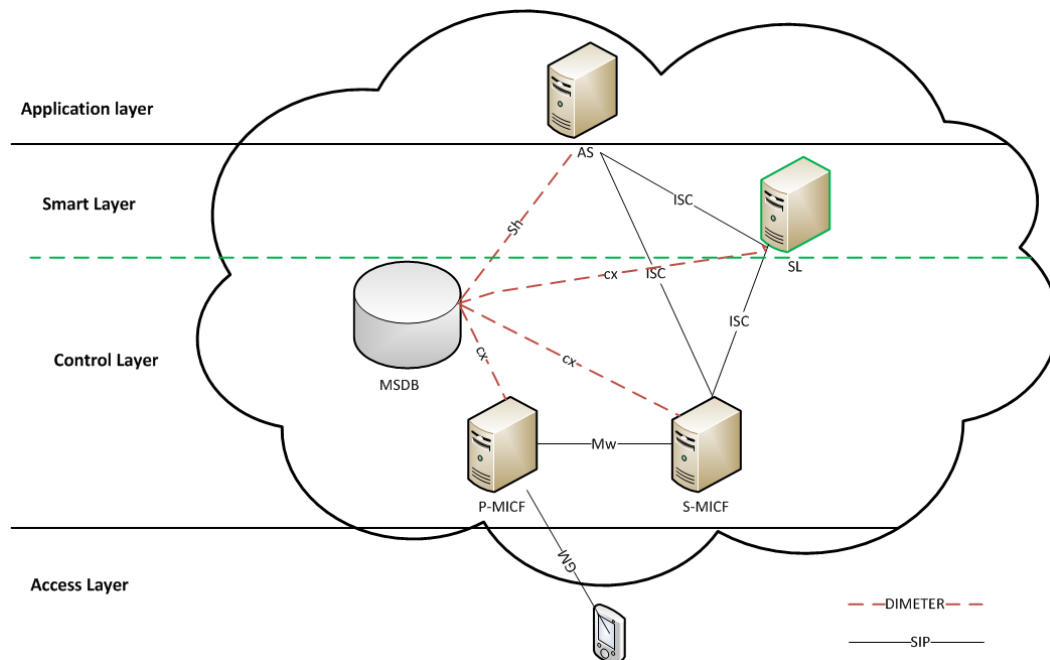


Figure 3.4 MICF Architecture.

Legend

P-MICF: Proxy Medical Information Control Function
S-MICF: Serving Medical Information Control Function
SL: Smart Layer
MSDB: Medical Subscriber Database
AL: SIP Application Layer server

3.6 Description of the Framework

This section discusses the framework entities and key functionalities. These entities can be classified into the following main categories:

- Access layer
- Session management and routing family (MICF) and database (MSDB)
- Smart layer
- Services (application server)

These four categories are described further below.

3.6.1 First layer: Access Layer

This layer consists of different types of users (medical staff from different specialisations such as clinicians, nurses, consultants, surgeons, healthcare students and patients) who need to access the library from their devices. These users range from GPs (diagnostic tools, PDAs) to home users, on PCs and mobile phone devices. This layer was created in order to achieve the design objective of allowing multiple methods of accessing information and flexibility of independent access.

3.6.2 Second layer: Control layer

The control layer is a SIP server which hosts the core network Medical Information Control Function (MICF). This layer separates the application layer and the entry users' devices. It is responsible for connectivity, and for establishing and handling sessions. This layer created in order to provide a scalable system to a large number of subscribers. It is comprised of three separate entities described below.

3.6.2.1 Proxy Medical Information Control Function P-MICF

The proxy is the first contact for the users, and the P-MICF acts as a gateway for the architecture terminal, which means that all messages sent by the UE will traverse the P-MICF. The P-MICF forwards SIP requests and responses in the appropriate direction, to

either the S-MICF or the UE. Moreover, the P-MICF provides a security function, which is achieved during registration, when the UE and P-MICF negotiate the user ID. Once the registration request is initiated, the P-CSCF confirms the identity of the user to the other units in the system. In addition to the security function, the P-MICF performs the following main functionalities:

- Receives the user requests whether it is a registration or an INVITE messages
- Forwards the registration requests received from the user to the S-MICF.
- Handles the SIP response from the S-MICF to the UE.

3.6.2.2 Server Medical Information Control Function S-MICF

The SIP server is the main server in the system it performs session control and maintains the session status. This unit receives messages sent by the UE via the P-MICF, whereupon it authenticates the user by downloading the user profile from the MSDB. This process is carried out by performing the user authentication request to the MSDB over CX to the database (see Chapter 4 for further details). Once that is done the S-MICF will be able to accept or reject the registration by sending 104 unauthorized message to the UE or 200 OK response. After this process the user is able to commence the session see chapter4 (INVITE process). In addition to the above, The S-MICF performs the following functions:

- Downloads the authentication vectors for the users by querying the HSDB during registration.
- Decides whether a user is eligible to access a particular service or not and whether the user's SIP messages will be forwarded to the application servers.
- Forwards the SIP messages for the application services.
- Decides whether to invoke the smart layer.
- Forwards the SIP messages initiated by the expert user to the smart layer.

3..3.2.3 Medical Subscriber Databases MSDB

The MSDB is the main system database for all users. It contains user-related information such as the user identity offered for registration and authorisation purposes, location information, security information and the parameters used to set up sessions, including parameters such as user authentication, roaming authorisation and allocated server name (Camarillo and García-Martín, 2009). This information is downloaded from the database upon receipt of requests related to authentication from other units. For this purpose, the HSDB provides a diameter protocol by which to communicate with other units, as illustrated in figure 3.5. The MSDB protocols will be described later in this chapter.

The authentication that the MSDB provides will confirm users' identification when the users are signed into the system; usually that can be through identifying users' names and passwords when they access the system. The authorisation determines whether the user eligible to access a specific resource or not based on the user's identity after logging into the system. It also determines what types or qualities of activities, resources or services a user is permitted. Once it has authenticated a user, it will send the response back to the unit (Curphey et al. 2002).

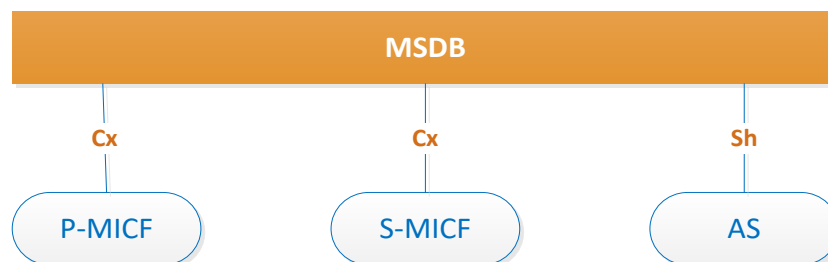


Figure3. 5 The MSDB Interfaces

3.6.3 Third layer: The smart Layer

This is the third layer in the framework located within the control layer and created for the purpose of customizing the search results. A description of the smart layer is covered in chapter 5.

3.6.4 Forth layer: Application Layer

This layer is an SIP entity consisting of different types of library services provided by medical libraries, such as: database search services, e-books, journals, library catalogues and other services. This layer has two types of interface; the first interface is based on the SIP protocol to contact the S-MICF, while the other is based on a diameter protocol to contact the MSDB. This layer created in order to achieve the objective of allowing new applications and services to be introduced without any major changes being undertaken.

From the description above, it is clear that the requirements of the framework satisfied the design objectives mentioned earlier in this chapter, by achieving them the research adapts the layered approached based on the layered module IMS and introduce the mentioned four - horizontal-layered architecture.

3.7 Overview of the Protocols and Interfaces in the Architecture

This section clarifies how the previously discussed components are all interlinked and the types of protocols and interfaces used to communicate between each entity in the architecture. As mentioned earlier, the framework is based on SIP and takes advantage of the IMS layered architecture. Therefore, the interfaces and protocols have been replicated from the already tested IMS architecture and integrated within the overall structure. However, the interfaces are standard interfaces in communication system. Figure 3.6 below illustrates the IMS reference points.

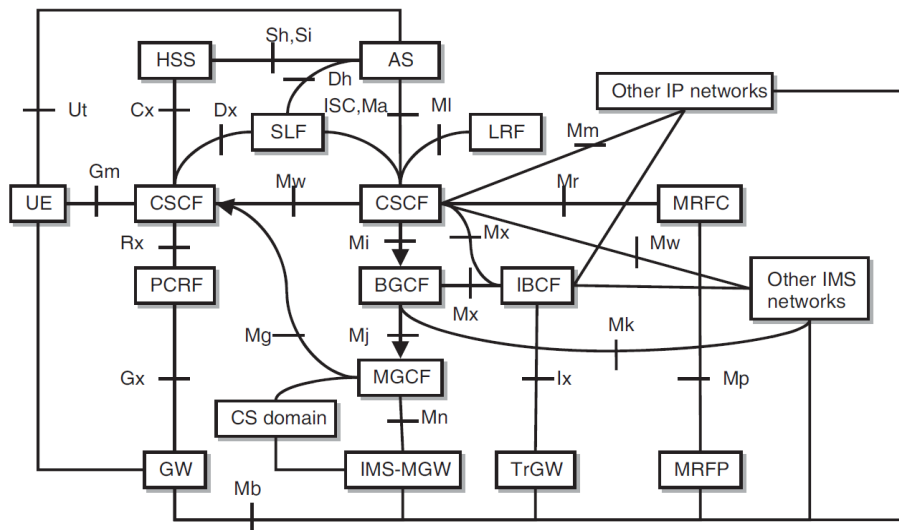


Figure 3. 6 IMS Reference points

3.7.1 The protocols

Two types of protocol are presented in the architecture: the first is used for the purpose of controlling the sessions, and is known as the Session Initiation Protocol (SIP), and the second is an AAA protocol used to perform the functions of authentication, authorisation and accounting. An overview of protocols and interfaces is presented in figure 3.7.

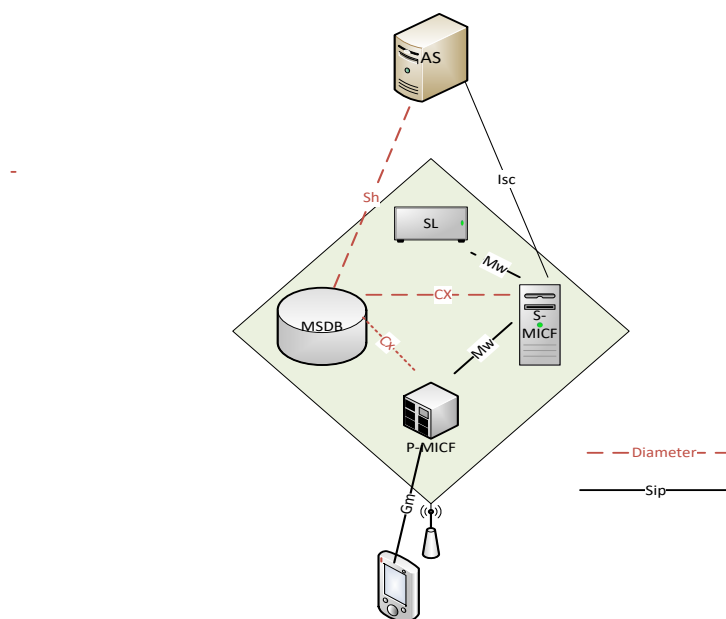


Figure3. 7 Overview of Protocols and Interfaces

3.7.1.1 SIP Protocol

Session Initiation Protocol (SIP) is an application-layer protocol used to create, modify, and terminate sessions over Internet Protocol (IP) networks. It can be also used to invite participants to the existing session such as a conference (Simon, 2005).

SIP message contains number of headers field that appear in every message (for further details, see section 4.2.5) as well as a number of request methods that indicate to the purpose of the request. These methods are presented in table 3.1. In contrast, SIP contains six categories of response codes which indicate the answer; these methods are defined in table 3.2.

Table3. 1 SIP Request Methods

Method name	Meaning
ACK	Acknowledges the establishment of a session
BYE	Terminates a session
CANCEL	Cancels a pending request
INFO	Transports PSTN telephony signalling
INVITE	Establishes a session
NOTIFY	Notifies the user agent about a particular event
OPTIONS	Queries a server about its capabilities
PRACK	Acknowledges the reception of a provisional response
PUBLISH	Uploads information to a server
REGISTER	Maps a public URI with the current location of the user
SUBSCRIBE	Requests to be notified about a particular event
UPDATE	Modifies some characteristics of a session
MESSAGE	Carries an instant message
REFER	Instructs a server to send a request

Table 3. 2 SIP f Response Codes

Status code	Meaning
100–199	Provisional (also called informational)
200–299	Success
300–399	Redirection
400–499	Client error
500–599	Server error
600–699	Global failure

3.7.1.2 AAA protocol

The AAA is used to refer to the authentication, authorisation and accounting offered via a number of interfaces in the system. First is the CX interface which connects the MSDB with the S-MICF. The second interface is the SH between the HSMD and the AS. This protocol provides the required protection in terms of accessing the system.

3.7.1.2 .1 Authentication

“Authentication is the act of verifying a claimed identity in the form of a pre-existing label from a mutually known name space, as the originator of a message (message authentication) or as the end point of a channel (entity authentication).”

Authentication involves validating the end user’s identity in order to allow them to access the system and establish a session. This process requires the end user to submit unique information to the system, which may include a username-password combination or a secret key, which is similar to the way in which access is gained to an email address. When the system receives this information, the AAA compares the user-supplied authentication data with the user-associated data stored in the Medical Subscriber Database (MSDB). If the data

match, the user will be able to access to the system; otherwise, the system will reject the user request (B. Abooba, et al 2001).

3.7.1.2.2 Authorisation

“Authorisation the act of determining if a particular right, such as access to some resource, can be granted to the presenter of a particular credential.”

The authorisation defines which services an end user is allowed to access upon receiving authorised access to the system. This might include providing an IP address or invoking a filter to determine which applications are supported. Authentication and authorisation are usually performed together in an AAA-managed environment. (B.Abooba,et al,2001).

3.7.1.2.3 Accounting

“Accounting the act of collecting information on resource usage for the purpose of trend analysis, auditing, billing, or cost allocation”

Accounting, as the final “A” in AAA, provides the method by which information is gathered on an end user's resource consumption, which can then be processed for billing, auditing and capacity-planning purposes, and also for the purpose of handling abuse, in order to monitor and act against malicious users (B.Abooba,et al,2001)

3.7.2 The Interfaces

The interfaces used to connect the framework entities in order to allow the messages to transfer from unit to the others, there are five main interfaces used in the framework, these interfaces are already existing in a number of communication systems such as IMS, ITU.

3.7.2.1 Gm interface

Gm is the first interface used to exchange the SIP messages between the UE and the system. The procedure in the Gm interface divided into two main categories registration, and transactions:

- 1- **Registration process:** the UE uses the Gm interface to send the registration request to the P-MICF. During the registration the UE can exchange some information related to parameters for AS, as well as information regarding authentication.
- 2- **Transaction process:** the transaction process means that all the INVITE requests that user send to the system after the registration will be transformed to the proxy P-MICF via the Gm interface and all responses (e.g., 200 OK) back to the UE in the same way over the Gm interface.

3.7.2.2 Cx interface

Cx is a diameter based interface used to link the MSDB with the S-MICF during registration. When a user signs into the system, the S-MICE communicate with the MSDB to authenticate the user in order to allow them access to the system. User related data and service known as user profile are permanently stored in the MSDB. These centralized data need to be utilised by the S-MICF when the user registers or receives sessions. Therefore, there has to be an interface between the MSDB and the S-MICF.

In order to perform these functions, a number of command requests and answers for the CX interface are utilised, as defined in table 3.3.

Table 3. 3 Cx Interface Command Name and Code

Command Name	Abbreviation	Command code
User-Authorization-Request	UAR	300
User-Authorization-Answer	UAA	300
Server-Assignment-Request	SAR	301
Server-Assignment-Answer	SAA	301
Location-Info-Request	LIR	302
Location-Info-Answer	LIA	302
Multimedia-Auth-Request	MAR	303
Multimedia-Auth-Answer	MAA	303
Registration-Termination-Request	RTR	304
Registration-Termination-Answer	RTA	304
Push-Profile-Request	PPR	305
Push-Profile-Answer	PPA	305

3.7.2.3 Mw Interface

Mw is an SIP based interface used to link the P-MICF, S-MICF and SL. The Mw interface is used during the following processes:

- 1- The P-MISF uses the Mw to forward the registration request that received from the UE to the S-MICF and the response from the S-MCF back to the P-MICF via the Mw interface.
- 2- The S-MICF uses the Mw interface to forward the request to the SL.
- 3- The S-MICF uses the Mw inform the UE and the P-MICF about the status of the request.

3.7.2.4 Isc Interface

ISC is an SIP based interface used to link the S-MICF with the application services. The S-MICF is capable of requesting any service by soliciting AS. The main role of the ISC interface is to provide service invocation and present SIP parameters to application (Edward, et al, 2008) Moreover, the ISC is used for the following purposes:

- 1- To forward the request from the S-MICF to the AS.
- 2- To notify the AS of the registration status.
- 3- To supply the AS with information in order to allow it to execute multiple services.

3.7.2.5 The Sh Interface

Sh is a diameter based interface linking the AS and the MSDB, and is used to access subscriber information. The SH provides data related to user identities which are stored in the MSDB; it also performs other functions such as downloading Application Server data from the MSDB or uploading Application Server data to the MSDB. These data can be in relation to the public user identity allocated to the user, registration status or the initial filter criteria (Camarillo and García-Martín, 2009). In addition, the SH interface is used to perform the following:

- 1- Provides a subscription service to change the data stored in the MSDB.
- 2- Provides notification services – upon any change in user data, the MSDB sends a notification to the AS (Miikka and Georg, 2009).

The Sh interface contains eight diameter command names by which to support its functionality, as shown in table 3.4.

Table 3. 4 Commands Defined by the Diameter Application for the Sh Interface

Command name	Abbreviation	Command-Code
User-Data-Request	UDR	306
User-Data-Answer	UDA	306
Profile-Update-Request	PUR	307
Profile-Update-Answer	PUA	307
Subscribe-Notifications-Request	SNR	308
Subscribe-Notifications-Answer	SNA	308
Push-Notifications-Request	PNR	309
Push-Notifications-Answer	PNA	309

3.8 User profile

The purpose of this section is to define the user profile and the way in which the S-MICF and the AS download the user profile from the MSDB. In addition, it will provide a brief explanation of how to update the user profile in the MSDB.

3.8.1 Overview of the user profile

Each user profile is comprised of a set of user-related information stored in the MSDB; each profile contains a significant amount of information related to a particular user.

User profiles contain at least one private user identity and single service profile. Moreover, each service profile contains a number of public user identity and initial filter criteria applicable to the user. Figure 3.8 depicts the general structure of a user profile, which is divided into four parts (Camarillo and García-Martín, 2009):

- 1- Public user identities. Each public identification contains a tag to indicate whether a public user identity is banned.
- 2- Core network service authorisation. This contains a subscribed media profile identifier which identifies the SDP parameters allocated to each user.
- 3- Initial filter criteria. These determine which SIP requests must visit which Application Servers so that a particular service can be provided.
- 4- Shared initial filter criteria. This is an optional filed supported by the server and the database (Camarillo and García-Martín, 2009).

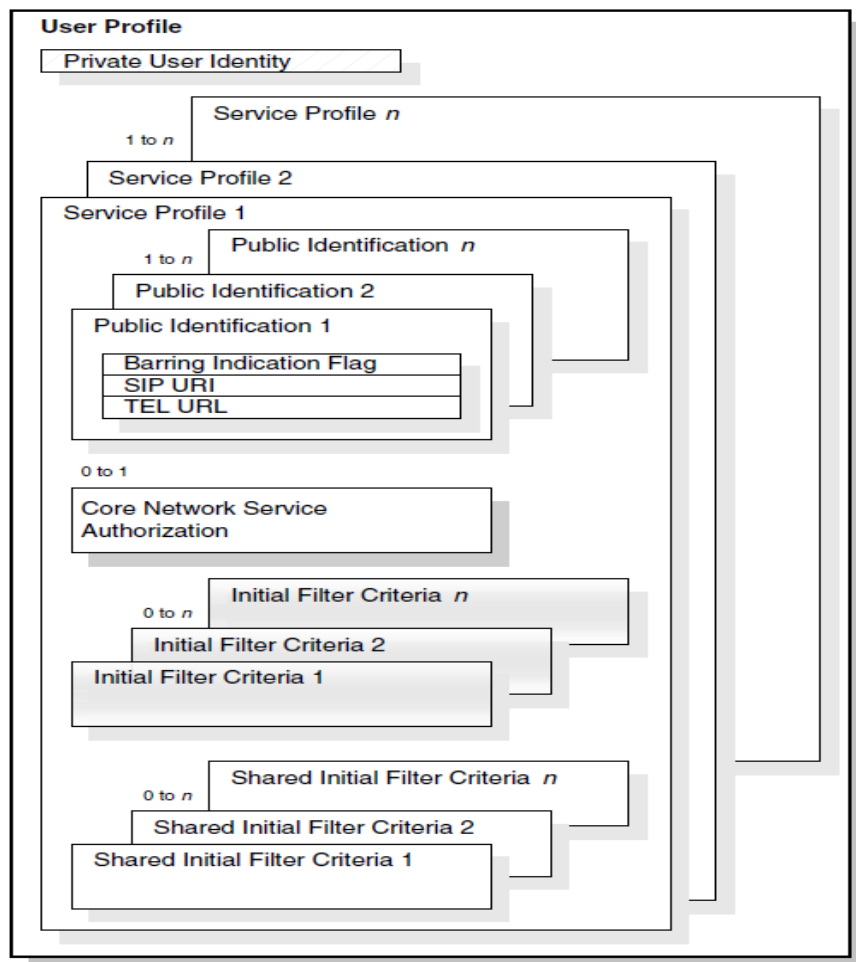


Figure3. 8 User Profile.

3.8.2 Public User Identity

The public user identity is utilised by the user when requesting communication with others. Each user is assigned one or more public user identity. A telephone number or an account similar to the email address (user@domain) can be used to identify the user, depending on the public user identities the users have (Camarillo and García-Martín, 2009).

This identity takes the form of either an SIP Uniform Resource Identifier URI as defined in RFC 3261 by J. Rosenberg, et al., (2002), or a Telephone Uniform Resource Locator TEL URI as defined in RFC 3966 (H. Schulzrinne, 2004).

Each public user identity has the following requirements:

- It takes the form of either SIP Uniform Resource Identifier (URI) or Telephone Uniform Resource Locator (tel URI) format.
- A public user identity needs to be registered before initiating the session.
- Public user identities are not authenticated by the network through registration. The UE cannot modify the stored public user identity.
- The public user identity is used for routing SIP requests.
- One user can have multiple identities (Miikka Poikselka and Georg Mayer, 2009).
- An identity can be associated with one user or a group of individuals (Al-begain et al., 2009).

Examples of public user identity is given.

Example of public User Identity	SIP:userA@example.com
--	------------------------------

3.8.3 Private user identity.

The private user identity is mainly used for authentication purposes; for example, it can be used for registration, authorisation, administration and accounting purposes. It is a unique identity defined by the home network operator. The private user identity is not using for routing SIP request and does not take the form of SIP URIs or TEL URIs. The requirement of this identity is listed below:

- It takes the format of a Network Access Identifier (NAI).
- It is contained in all registration messages sent by the UE.
- It is authenticated only during registration.
- It is permanently allocated to a user's subscription and remains valid throughout the subscription period.
- It is used to identify the user's data that are stored within the database.
- UE cannot modify any private user identity already stored.
- It must be obtained by the server and stored in the databases (MiikkaPoikselka and Georg Mayer, 2009).

Example of Private User Identity	userA@operator.com
---	---------------------------

Relationship between Private and Public User Identities.

According to the IMS, the home network operator is in charge of the assignment of private user identities and public user identities; other identities that are not defined by the operator may also exist (J. Rosenberg.et al., 2002). Moreover, both private user identities and public user identities allocated to subscribers are stored within the databases.

As mentioned earlier, user profiles consist of user subscription information such as private user identity and public user identity. This information is stored within general databases and downloaded to the server via a diameter CX interface. According to IMS, public user identities must be registered with a single S-CSCF on one occasion. Thereafter, the service profile is downloaded from the database HSS to the server S-CSCF. Each public user identity is associated with one service profile only, while each service profile is associated with one or more public user identity. Moreover, the service profile also contains the initial filter criteria, which provide a simple service logic comprising user / operator preferences that are of static nature (i.e., they are not changed on a frequent basis). It is possible to identify the public user identities of a user that are associated with a common service profile and have the same service configuration for each identity (J. Rosenberg.et al., 2002).

A public user identity may be shared by more than one private user identity within the same IMS subscription. Hence, a particular public user identity may be registered from multiple UEs associated with different private user identities and different contact addresses. If a public user identity is shared among the private user identities of a subscription, it is assumed that all private user identities in the IMS subscription share the public user identity (Al-begain et al., 2009).

Figure 3.9 depicts the relationship between a shared public user identity and a private user identity.

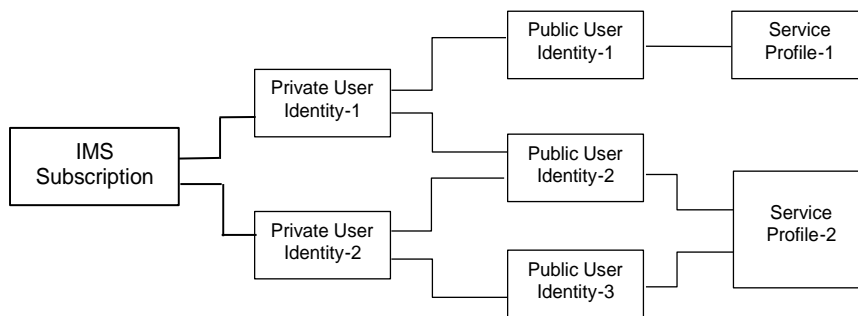


Figure 3. 9 Public and Private User Identities

3.8.4 Initial Filter Criteria.

Initial filter criteria are defined as a set of service point triggers (STPs) that describe when an incoming SIP message is routed to a particular AS. A user profile may hold both a user's initial filter criteria and a reference value to initial filter criteria, which are locally administered and stored in the server.

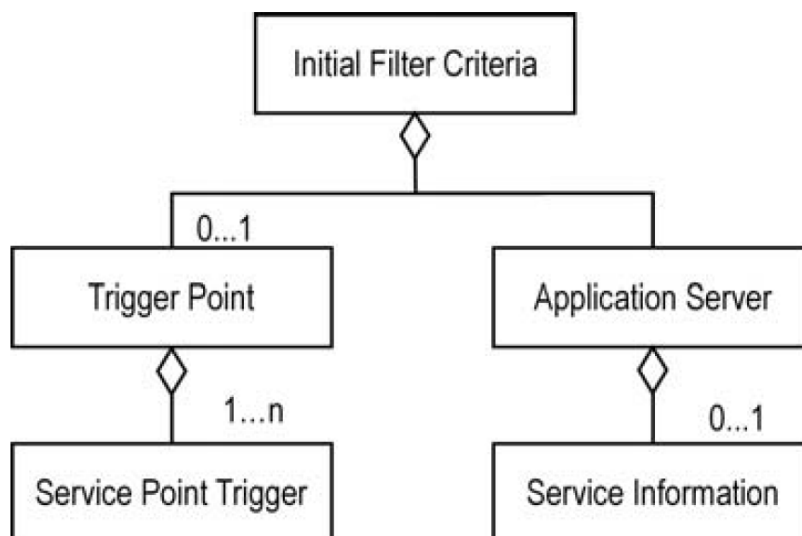


Figure 3. 10 Initial Filter Criteria

Figure 3.10 shows that Initial Filter Criteria are contain of either zero instances or one instance of a trigger point and one instance of an AS. Each initial filter criteria within the service profile has a unique priority value (integer) that is utilised in the server.

When multiple initial filter criteria are assigned, the server assesses them in numerical order; that is, an initial filter criteria with a higher priority number will be assessed after one with a smaller priority number.

The filter criteria are downloaded to the server when the server sends a request to the HSDB. Once the server has downloaded the user profile from the database, it will assess the filter criteria for the initial request, based on the following steps:

1. It ascertains that the public user identity is not barred and then proceeds.
 2. It checks whether the request originates or terminates.
 3. It selects the initial filter criteria for the session.
 4. It checks whether the request matches the initial filter criterion with the highest priority for the user by comparing the service profile with the public user identity that was used to send the request.
- If this request matches with the initial filter criterion, the server will forward the request to AS, check that it matches the following filter criterion of lower priority and apply the filter criteria on the SIP method received from the previously contacted AS.
 - If it does not match the highest priority initial filter criterion, the server will check that it matches the following filter criterion priorities until one does match (MiikkaPoikselka and Georg Mayer, 2009).

3.8.5 Trigger point

The trigger point contains a set of criteria or conditions that the session must meet in order to be routed to a specific application. These criteria for the trigger point are defined as a set of Service Point Triggers (SPTs). The service trigger point is linked through logical expressions (AND, OR, NOT), in order to specify the relationship between groups of SPTs and members of each group. The service trigger point contains a number of items, as shown in figure 3. 11.

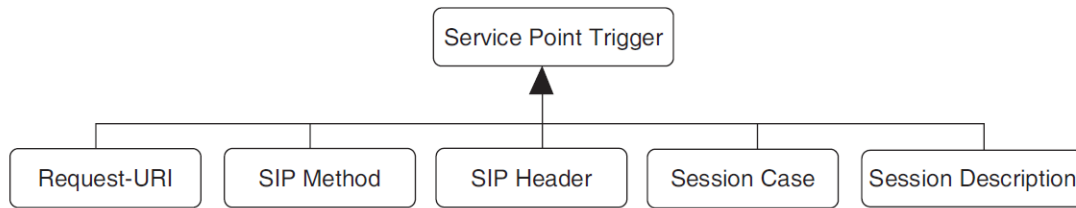


Figure 3. 11 Server Trigger Point Structure

- Request URI – identifies a resource that the request is addressed to (e.g., [User A@example.com](#)).
- SIP Method – indicates the type of request (e.g., INVITE or REGISTER).
- SIP Header – contains information related to the request. A Service Point Trigger can be based on the presence or absence of any SIP header or the content of an SIP header.
- Session Case – can be any one of four possible values – originating, terminating, originating unregistered or terminating unregistered – that specify whether the filter should be used by the server responsible for handing the service to the end user.
- Session Description – defines an SPT for the content of any SDP field within the body of an SIP method (MiikkaPoikselka and Georg Mayer, 2009).

3.8.6 User Profile Data Request and Update.

This section will describe how the server request the user profile from database, and how to update the user profile.

3.8.6.1 Data Requesting by the S-MICF

When the user registered with the system S-MICF can request the user data by sending a diameter Server-Assignment-Request to the MSDB for the purpose of informing the MSDB that the user is now registered in the S-MICF and to download

the user profile. The user data can be of a type defined over a CX interface. The MSDB returns the result and attaches the user profile by way of a diameter Server-Assignment-Answer. Figure3.12 depicts the process of this call flow.

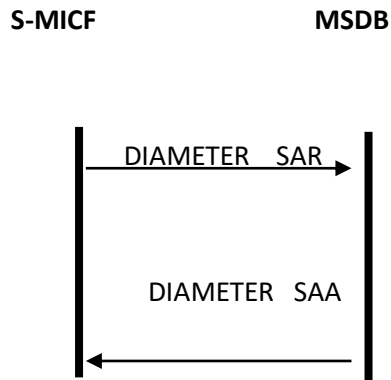


Figure3. 12 Process of S-MICF Data Request from MSDB

3.8.6.2 Data Requesting by the Application Server

The application server requests user related data from the MSDB by sending a diameter User-Data-Request (UDR) for a particular user. The user data can be of a type defined over the SH interface. The MSDB returns the requested type of data by way of a diameter User-Data-Answer (UDA) message. Figure 3.13 depicts the process of this call flow.

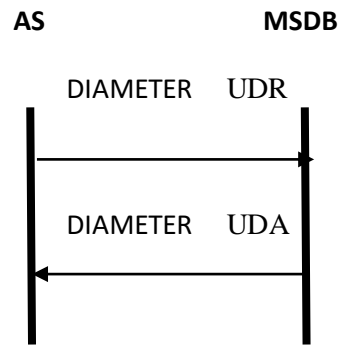


Figure3. 13 Process of Application Server Data Request to MSDB

3.8.6.3 User profile Update

Each user has a user profile associated with him/her, together with other accounting information. This profile may change with time and therefore information needs to be updated, for instance in relation to a user's subscription to a new service. An AS may modify user data and store them in the MSDB. In order to do this, a diameter Profile-Update-Request (PUR) is sent by the AS to update the user related data. Figure3.14 depicts the applicable call flow.

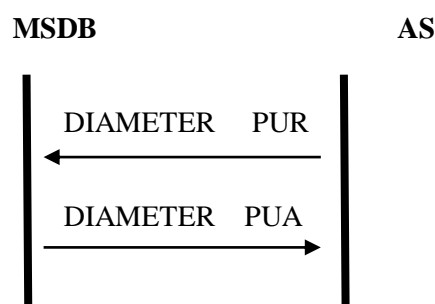


Figure3. 14 Process of User Profile Updates

The MSDB can also send messages to the S-MICF containing updated information associated with the user. This allows an operator to modify the data of a user profile or the accounting information and transfer it to the SIP server where the user is registered. This procedure is used by the database when a new service is available to the user and the MSDB requires new filter criteria to be added to the user profile. In addition, the MSDB may add a new public user identity. In this case, the MSDB will send a diameter Push-Profile-Request to the S-MICF. The server returns the requested type of data in a diameter Push-Profile-Answer (PPA) message. Figure 3.15 shows this call flow.

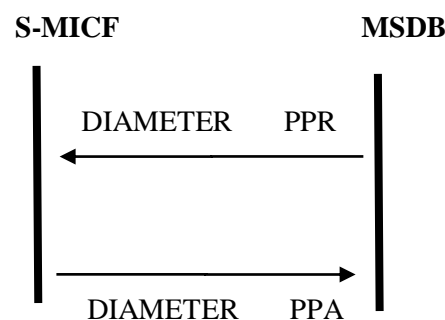


Figure3. 15 Data request by the MSDB

A user profile can be downloaded to the S-MICF when the user registers for the first time with the system, and also when the user send an INVITE messages.

3.9 Framework Contraption

The proposed framework is a novel framework that created for the medical information and no one has investigated enough in that area especially the framework inspired by the IMS. This framework is a novel by nature as it is a first framework based on our reading that inspired by the communication system IMS and managed to allow the user to use different access methods to retrieve the information. And to get to the results as fast as possible and

targeted the results based on each user profile. The designed framework has been implemented and evaluated in chapter 6. Moreover, the smart layer will be presented in chapter 5 is another contraption that adds value to the system.

3.10 Summary

In this chapter, a layered architectural framework is designed. The four main layers are defined and the functions of the framework entities are described. This layered model is proposed for use with medical information domain. The designed framework takes advantages of the existing IMS architecture and adapts the horizontal layers as well as the protocols and the interface that used to connect the frameworks entities to each other. The chapter has concluded with a details description of the user profile along with the user identity.

The next chapter will provide a detailed description of the signalling plan for the framework.

Chapter 4: Signalling Plan

The previous chapter introduced the framework and the tasks required for each functional entity. This chapter presents the signalling plan that applied to investigate the way that the framework is handling the transaction and the process of the user request and answer. Chapter 3 introduced number of protocols and interfaces that used to connect the framework entities with each other, this chapter first, described how these protocols can be extended for use in the framework and how that protocols fit into the framework, followed by a call flow explanation of the registration process and the session set up. That will help the reader to understand how the framework handles the transaction and provides the service to the user. Finally, the chapter will conclude with a two scenarios of two types of searches (the general search and the expert search) to demonstrate the data flow plan of the request.

The signalling plan is the path that created to allow the messages to be exchange from node to node in the system, the most important component of the signalling plane is the protocol that performs the user session. As mentioned earlier in chapter 3 that SIP is the protocol that chosen to perform the sessions in our system due to its flexibility and descriptive power not only for performing sessions but also for other services such as multimedia streaming and web-services. Before starting describing the signalling plan or the call flow, a brief description for the SIP message format will be presented in this section, in order to give the reader a broad perspective of the contents of the message.

4.1 SIP Address

Session Initiation Protocol (SIP) is a signalling protocol used for initiating, maintaining, modifying and terminating real-time sessions that involve video, voice, messaging and other communications applications and services between two or more endpoints on IP networks RFC3261 (J. Rosenberg.et al., 2002).

In SIP environment the users will be identified by SIP Uniform Resource Identifiers (URIs) format, this format of a SIP URI required a use name and password, which is similar to the e-mail, address for instance: *userA@domain.com*

4.2 Message Format

Before introducing the SIP message description format, it is important to clarify that the SIP format is a standard format that appears in any SIP message. Figure 4.1 shows the format of SIP messages as described in RFC3261 (J. Rosenberg.et al., 2002). It consists of a start line followed by a set of header fields, and an empty line indicating the end of the header fields, and an optional message-body. Detailed information about each one will describe in this chapter.

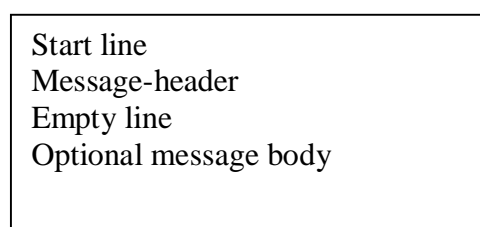


Figure 4. 1 SIP Message Format

4.2.1 Start line of SIP Message Request and Response.

At any SIP transaction, the message contains a request from the UE and the response from the server. The request will invoke a particular function on the server and at least one response from the server. In the registration process for example (that will be described later on this chapter), the transaction begins with the UE sending a Register request addressed to the server. This message contains a start line for both the request and response.

The first line in the SIP request message called request line, which refers to the purpose of the request (REGISTER, INVITE...). This line consists of the SIP message type, SIP protocol version and requests the URI. The SIP message type or SIP methods are standard SIP methods existing in a communication system, which can be used to indicate the purpose of the request (Camarillo & García-Martín, 2009). SIP methods are illustrated in table 4.1. The start line format for the request is illustrated in the following example

REGISTER sip:userA@example.com SIP/2.0

To clarify the above example, the phrase REGISTER is a SIP method that indicates the action that the server should do, the Request-URI shows that this request is intended for user A and is similar to the email address that consists of user name and domain name, and finally the SIP protocol version SIP/2.0. In addition to the start line, the SIP request contains a number of header fields that provide additional information about a message; it will be described later on in this section.

Moving to the start line of the SIP response message, which is called the status line as it refers to the status of the response, it is a response message generated by the server to reply to the user request. This line contains the SIP version, a response code that is divided into 6 classes

(as shown in Table 4.2) and finally the response (Camarillo & García-Martín, 2009). The example below shows the format of the status line.

SIP/2.0 200 OK

The example above shows the protocol version that is always set to SIP/2.0 (Camarillo & García-Martín, 2009) followed by the response code 200 OK success.

Table 4. 1 SIP Request Methods

Method name	Meaning
ACK	Acknowledges the establishment of a session
BYE	Terminates a session
CANCEL	Cancels a pending request
INFO	Transports PSTN telephony signalling
INVITE	Establishes a session
NOTIFY	Notifies the user agent about a particular event
OPTIONS	Queries a server about its capabilities
PRACK	Acknowledges the reception of a provisional response
PUBLISH	Uploads information to a server
REGISTER	Maps a public URI with the current location of the user
SUBSCRIBE	Requests to be notified about a particular event
UPDATE	Modifies some characteristics of a session
MESSAGE	Carries an instant message
REFER	Instructs a server to send a request

Table 4. 2 SIP Response Code.

Status code	Meaning
100–199	Provisional (also called informational)
200–299	Success
300–399	Redirection
400–499	Client error
500–599	Server error
600–699	Global failure

4.2.2 Header Fields

As mentioned earlier, the start line will be followed by a number of header fields to provide additional information about a message. A header field consists of the header field's name, a colon, and the header field's value, as shown in the following example:

To: user A <sip: userA@example.com>;tag=1234

The (To) header field above contains a name (user A) and an URI (sip: userA@example.com), followed by a tag parameter. The header (To) in the example is one of six mandatory fields that appeared in every SIP message; each one describes a particular piece of information in the message:

- **To:** this header field specifies the address of the user or the resource that is the target in this request. The final destination of the To header field will be selected by the user via a human interface, perhaps entering the URI manually or selecting it from some sort

of address book. the following is an example of the To header field RFC3261 (J. Rosenberg.et al., 2002)

```
To: Carol <sip:carol@chicago.com>
```

- **From:** the (From) header field contains the URI of the originator of the request possibly the uses'address. Like the To header field, it contains a URI and display name. The following is an example of the From header field RFC3261 (J. Rosenberg.et al., 2002)

```
From: "Bob" <sip:bob@biloxi.com>
```

- **Cseq:** the Cseq header field serves as a method to identify and order transactions. It consists of a sequence number and a method name. The method must match the requests and responses. Example of this field is given by (J. Rosenberg.et al., 2002).

```
CSeq: 4711 INVITE
```

- **Call-ID:** the Call-ID provides a unique identifier for a SIP message exchange. It must be the same for all requests and responses sent by the user. An example of the Call-ID header field as defined in the RFC3261 (J. Rosenberg.et al., 2002) looks like the following :

```
Call-ID: f81d4fae-7dec-11d0-a765-00a0c91e6bf6@foo.bar.com
```

- **Max Forwards:** the Max-Forwards header field used to avoid the number of loops. Every proxy that handles a request decrements its value by one, if the Max-Forwards value reaches zero before the request reaches its destination, the request will be rejected with a 483(Too Many Hops) error response.RFC3261 (J. Rosenberg.et al., 2002).

Example: Max-Forwards: 70

- **Via:** The Via header field indicates to the transport used for the transaction and recognizes the location of the response. This header field keeps track of all the proxies

a request has traversed. The Via header field value contains a branch parameter used by both the client and the server. This parameter is used to identify the transaction created by that request. An example of the header Via is given by and shown below (J. Rosenberg.et al., 2002)

```
Via: SIP/2.0/UDP pc33.atlanta.com; branch=z9hG4bKhjhs8ass877.
```

Note that the From field of the response MUST equal the From header field of the request. The Call-ID header field of the response MUST equal the Call-ID header field of the request. The CSeq header field of the response MUST equal the CSeq field of the request. The Via header field values in the response MUST equal the Via header field values in the request and MUST maintain the same ordering SIP RFC 3261(J. Rosenberg.et al.2002).

To clarify the header fields to the reader. An example of the header fields for the REGISTER message is shown in figure 4.2 for the request and figure 4.3 for the respond. Both figures are selected from the already tested IMS SIP REGISTER message .RFC 3261(J. Rosenberg.et al.2002).

```
REGISTER sip:domain.com SIP/2.0
Via: SIP/2.0/UDP 192.0.0.1:5060;branch=z9hG4bKna43f
Max-Forwards: 70
To: <sip:Alice.Smith@domain.com>
From: <sip:Alice@pda.com>;tag=453448
Call-ID: 843528637684230@998sdasdsfsgt
Cseq: 1 REGISTER
Contact: <sip:alice@pda.com>
Expires: 7200
Content-Length: 0
```

Figure 4. 2 Header Field for SIP REGISTER Message

```
SIP/2.0 200 OK
Via: SIP/2.0/UDP 192.0.0.1:5060;branch=z9hG4bKna43f
    ;received=192.0.0.1
To: <sip:Alice.Smith@domain.com>;tag=54262
From: <sip:Alice@pda.com>;tag=453448
Call-ID: 843528637684230@998sdasdsfgt
Cseq: 1 REGISTER
Contact: <sip:alice@pda.com>
Expires: 7200
Content-Length: 0
```

Figure 4. 3 Header Field for SIP 200 OK

4.2.3 Message Body

The SIP message body carries session descriptions to provide additional and richer information to describe the session. These session descriptions are formatted using session description protocol SDP. An SDP session description consists of a number of lines; each line has a reference to the type = value. The type is one character and the value is structured text whose format depends on <type>.

An SDP is created using a standard format that consists of a session-level section start by (v) followed by zero, or more media-level sections that start by (m). The description has to be in the same order starting with the session level then the media level. Table 4.3 below shows the format lines of the SDP. (Camarillo & García-Martín, 2009). The example of SDP will be shown later in section 4.3.2.

Table 4. 3 SDP Type and Value

Type	Value
V	Protocol version
B	Bandwidth information
O	Owner of the session and session identifier
Z	Time zone adjustments
S	Name of the session
K	Encryption key
I	Information about the session
A	Attribute lines
U	URL containing a description of the session
T	Time when the session is active
E	Email address to obtain information about the session
T	Times when the session will be repeated
P	Phone number to obtain information about the session
M	Media line
C	Connection information
I	Information about the media line

From the description above, all elements of the SIP message are clarified. The next section will present the registration procedure.

4.3 The Registration

The registration is a required procedure for all users in which they request authorization to use the services. It requires sending a request, This request would contain, an identity to be

registered and a home domain name for example userA@example.com which is similar to the email address to be send to the server which is known as a registrar, this registrar acts as the front end to the location service for a domain, processing the request based on the contents of requests. A proxy server that is responsible for routing requests for that domain RFC3261 then typically consults this location service. An illustration of the overall registration process is shown in figure 4. 4.

The registration procedure contains two phases: first phase challenge send by the server to the user, and the second phase is UE responds to the challenge and completes the registration. The registration process will be defined in the following steps.

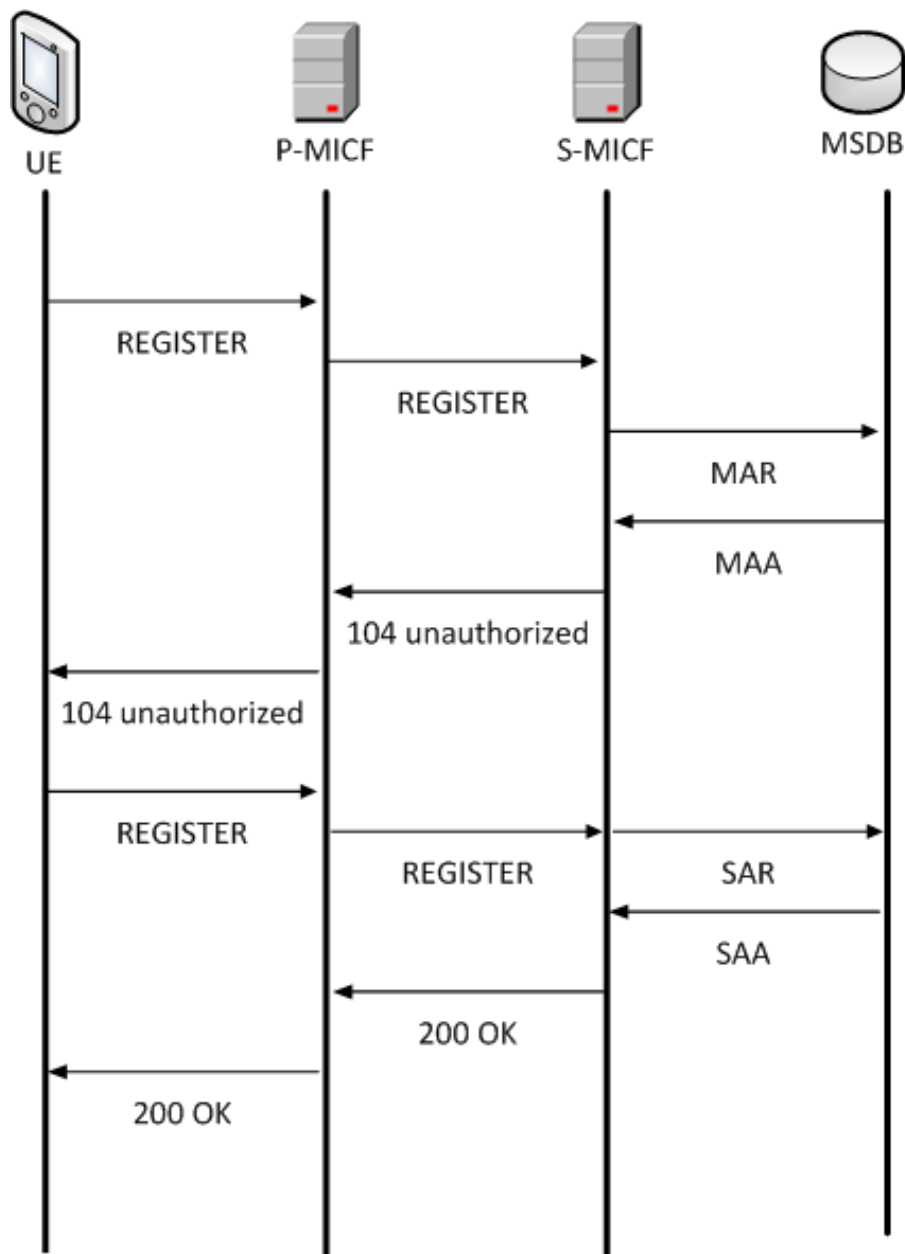


Figure 4. 4 Registration Procedure

4.3.1 From the UE to the S-MICF

The registration starts when the UE sends the SIP request to the first hop in the system, which is the proxy P-MICF via the Gm interface. This request would contain the user identity and a home domain name similar to the email address. When the P-MICF receives the registration request it forwards the request to the S-MICF via the Mw interface to confirm that the user is

here and needs to register, and to check the authorization information. This information will help the P-MICF to make a decision whether or not to continue with the registration.

4.3.2 From the S-MICF to the MSDB

Having done the above step, the P-MICF forwards the SIP register request to the server S-MICF via the Mw interface. When the server receives the SIP register request from the P-MICF it will recognize that the user is not authorised and has to be authenticated (see 4. 4).

Therefore, the S-MICF will send Diameter Multimedia-Authentication-Request (MAR) message to the MSDB for downloading user authentication data. The MSDB responds with a Diameter Multimedia-Authentication-Answer (MAA) message with one or more authentication vectors. When the S-MICF retrieves the authentication data from the MSDB it will challenge the user with a 401 unauthorized response that UE must answer it, this challenge will be delivered to the UE via the proxy over GM interface. Afterwards, the UE calculate the response to the challenge and send new SIP register request response.

Note that the second register request will include the same registration-related information and will be routed in exactly the same way as the initial register request. If the authentication is successful, the S-MICF sends a Diameter Server-Assignment-Request (SAR) over the Cx interface to inform the MSDB that the user is registered to the S-MICF is now able to retrieve the data and download the user profile that contains the Public User Identities allocated for authentication of the Private User Identity. It also indicates to the S-MICF in which of these Public User Identities are automatically registered. Additionally, the user profile also contains the initial filter criteria, which is the collection of triggers that determine which Application Server will be providing the service for the UE. After receiving the Diameter Server-Assignment-Answer (SAA) from the MSDB, the S-MICF will update the registration

information in the MSDB, and download the user profile. Moreover, the S-MICF stores the UE uniform resource indicator (URI) in the MSDB. The server now becomes aware of the address of UE, therefore all initial requests (INVITE requests) that are sent by a particular UE will be recognised by the S-MICF. Based on the initial filter criteria received from the MSDB, the S-MICF will inform any Application Servers (ASs) that are interested in UE's registration state. Afterwards, the server will forward the authorization answer and accept the registration with a 200 OK response to the P-MICF.

4.3.3 From the P-MICF to the UE

Once the P-MICF has received the 200OK response, it will forward the response back to the UE over the Gm interface to inform the UE that the registration is successfully completed.

As mentioned earlier in section 4.1 that the SIP message contains header fields which provide an important information to describe the session, figure 4.5 shows an example of the message header field of the SIP registration request followed by figure 4.6 the header field of the SIP registration response as described on FRC3216 Rosenberg.et al.(2002).

```
REGISTER sip:userA.med.library.com SIP/2.0  
  
Via: SIP/2.0/UDP . 192.0.0.1:5060;branch=z9hG4bKna43f  
  
Max-Forwards: 70  
  
To: <sip: userA@ med.example.test>  
  
From: <sip: userA@example.com >;tag=453448  
  
Call-ID: 843528637684230@998sdasdsfgt  
  
CSeq: 1 REGISTER  
  
Contact: <sip: userA@ med.example.test>  
  
Expires: 7200  
  
Content-Length: 0
```

Figure 4. 5 Register Header Field

SIP/2.0 200 OK

Via: SIP/2.0/ UDP 192.0.0.1:5060;branch=z9hG4bKna43f;received=192.0.0.1

To: < sip: userA@ med.example.test>;tag=54262

From: <sips: userA.med..example.com>;tag=453448

CSeq: 1 REGISTER

Contact: <sips: userA.med..example.com>

Expires: 7200

Content-Length: 0

Figure 4. 6 SIP 200OK Field

Note, that the message format in both figures 5.6 are not a complete REGISTER request; there are some headers and parameters are not included. It only includes the information required to explain the registration procedures in this section. This format shows that the header From contains the public user identity to identify the user who is performing the registration. Also within the via header of the request the UE put an IP address, to ensure that all responses to this request will be routed back to the UE. The To header field specifies the recipient of the request

4.4 Authentication

4.4.1 Overview

Authentication is the act of confirming that the communicating entity is the one claimed. It contains a request from the server, this request called a challenge in communication systems, and response from the user to answer the challenge; this is the way that most communication systems perform the authentication. In this study the authentication procedure is based on existing (IMS) authentication process that shown in figure 4.7 below (Sharma, G et al, 2011).

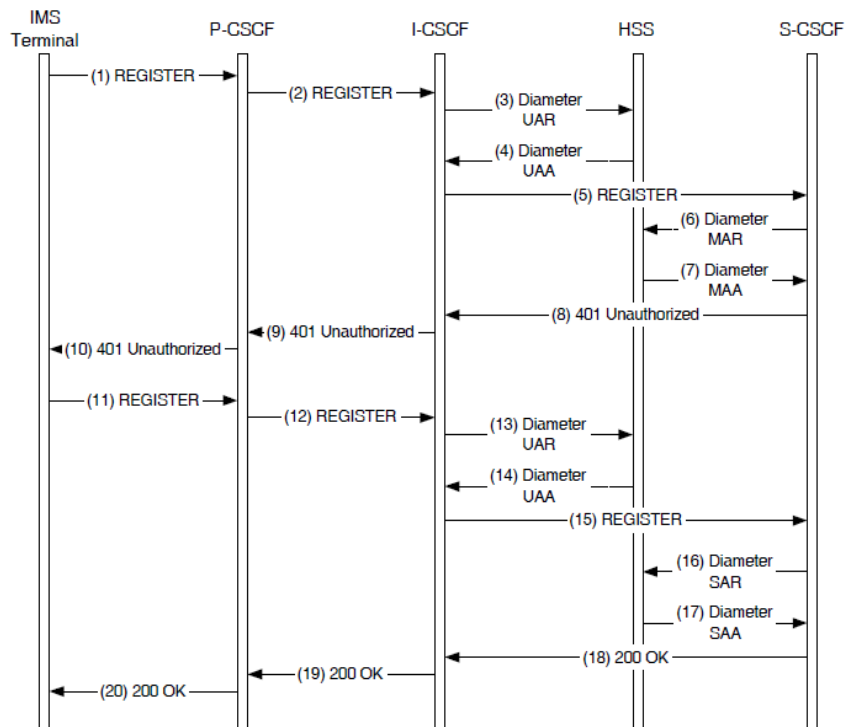


Figure 4. 7 IMS Authentication

According to the IMS the authentication required a unique information to be downloaded to the server, this information is named as Authentication Vectors (AVs). The authentication vectors contain a number of parameters which provide the authentication data that enable the server to authenticate the user. These parameters are available in the database, as that there is no direct connection between the database and the UE the Server S-MICF will be responsible of performing the authentication and downloading the AV's from the database over the Cx interface during the registration.

In order to authenticate the UE, the user will send his private user identity in the initial register request. This private user identity is stored within MSDB and is only used for authentication and registration procedures. When the registration reaches the server, it needs to take care that

the user is authenticated. Then it will download the AV from the database in order to allow the user to use the services.

The authentication vectors been defined in RFC3310 by A. Niemi., et all (2002) and ordered based on sequence number. Each authentication vector consists of the following parameters.

- AUTN : Authentication Token. This parameter authenticates the server to the client.
- CK : Cipher Key. An AKA session key for encryption.
- IK : Integrity Key. An AKA session key for integrity check.
- XRES: Expected Authentication Response. In a successful authentication this is equal to RES.
- RAND: Random Challenge. Generated by the home network's Authentication Centre (AUC). Using the shared secret and a sequence number SQN.

As mentioned earlier that authentication procedure is based on challenge and response between the server and the UE to perform the authentication. The following section will briefly describe the IMS authentication processes as described on RFC 3310 (A. Niemi., et al. 2002).

1. The authentication centre of the home network produces an authentication vector Av. Each AV contains the RAND, AUTN, XRES, IK, and CK. Moreover, it will be stored in the database.
2. The server will download the authentication vectors from the database over MAR/MAA message, the server may downloads more than one authentication vectors per a request.
3. The authentication challenge request (401 unauthorized) is created by the server, which contains the random challenge RAND, and the network authenticator token AUTN.
4. The authentication request is delivered to the user.

5. The user calculates the response and then produces an authentication response RES, using the shared secret K and the random challenge RAND.
6. The server sends a successful authentication 200 OK.

The above steps summarized the basic authentication procedure in IMS system. This study follows the same procedure to authenticate the user.

When the server S-MICF receives the registration request from the proxy P-MICF, the S-MICF needs to set the registration information at the MSDB for the initial request and to retrieve the authentication information associated to the user. The following steps will give a clear description of the authentication process. A description of the overall authentication process is showing in figure4. 8.

4.4.2 Authentication of the Initial Register Request

The authentication procedure is one of the main functions that the server S-MICF is responsible for, this process starts since the UE sends the first initial registration request. This initial request includes the user name field (private user identity) for example (userA@example.com). This private user identity is stored in the database MSDB and is used only for authentication and registration process. In addition to the realm and URI fields, and finally the response and nonce field.

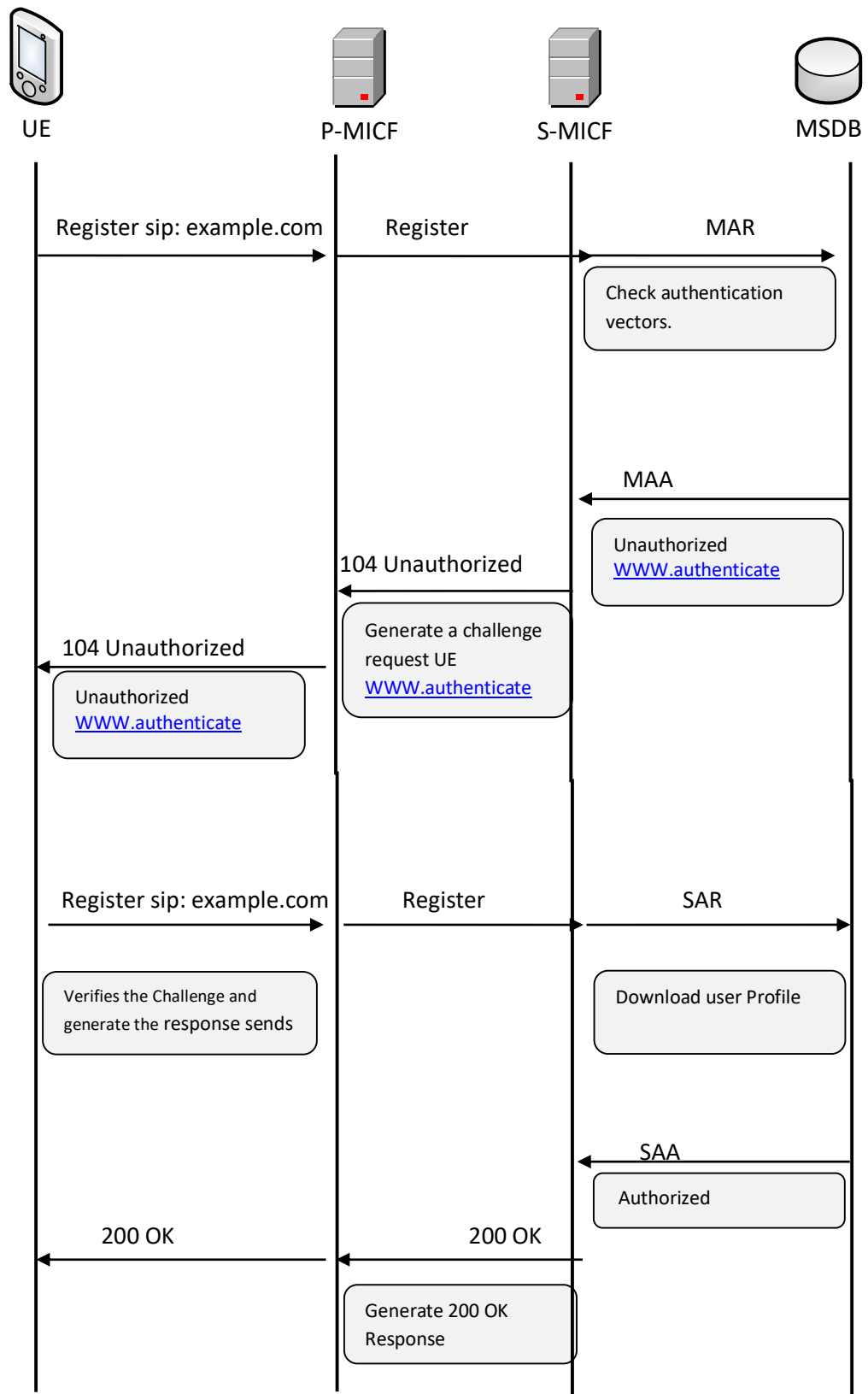


Figure 4. 8 Call Flow for the Authentication

Once the S-MICF received, the request will refer to the MSDB and download the following authentication vectors RAND, XRES, CK, IK and AUTN that are mentioned before. The S-MICF contacts the MSDB by performing a diameter Multimedia-Authentication-Request MAR over the Cx interface for two purposes: one is to download the authentication vectors, the second to register the URI associated to the UE in case if the user has not registered yet.

Based on the authentication vectors (AVs) that received on a diameter Multimedia-Authentication-Answer MAA that the S-MICF received from the MSDB, the S-MICF will authenticate the user. If the verification is successful, the S-MICF will treat the user authentication and will perform the SIP registration procedure (see section 4.2). Figure 4.8 summarizes the authenticating process for the user request.

Now the initial registration request header sent by the user will look as follows:

```
REGISTER sip:home1.fr SIP/2.0
Authorization: Digest username="Bob private@home1.fr",
realm="home1.fr",
nonce="",
uri="sip:home1.fr",
response=""
```

This above example contains a number of parameters that refer to the authentication the meaning of the values of the headers is described by (Miikka and Georg, 2009) as follows:

- The username field which shows the private user identity that will be used by the server and the database to identify the user and find the corresponding AVs.
- The realm header is a string to be displayed to users so they know which username and password to use. This string should contain at least the name of the host performing the authentication and might additionally indicate the collection of users who might have access.

- The nonce is a parameter, which is populated with the Base64 encoding of the concatenation of the AKA authentication challenge RAND, the AKA AUTN token, and optionally some server specific data.
- The URI associated to the user and set to the home domain of the user.
- The response: this left empty.

When the Proxy receives the request (initial registration), this request including the above parameters, the P-MICF cannot guarantee that the REGISTER request sent by the actual user. Therefore, the P-MICF adds the integrity-protected field with the value 'no' to the authorization header, before sending the request toward the S-MICF for the purpose of security, once the P-MICF add that field it will forward the request to the S-MICF.

The integrity-protected field is showing below as described by Miikka and Georg, (2009).

```
REGISTER sip: home1.fr SIP/2.0
Authorization: Digest username="tobias private@home1.fr",
realm="home1.fr",
nonce="",
uri="sip:home1.fr",
response="",
integrity-protected="no"
```

Further information on the header registration is explained in the RFC3310 by (A,Niemi., et al 2002).

4.4.3 S-MICF Challenges the UE

Upon information received in the MAA from the MSDB, the S-MICF will generate 401. This challenge will contain a number of fields as described in the IMS by Miikka and Georg, (2009).

These fields are:

- The nonce field it has the RAND and AUTN parameters, which were present in MAA. Both values are 32 bytes long and Base 64-encoded.
- The ik and ck extension fields it has the integrity and ciphering keys, as they were present in the MAA.

The user is currently not authenticated yet so, the S-MICF will reject the registration request and generate a 401 unauthorized challenge to be sent to the user. It will then send the 401 unauthorized to the P-MICF in order to send back to the UE. The authentication header of this challenge request as described by Miikka and Georg, (2009) will look as follows:

```
SIP/2.0 401 Unauthorized
WWW-Authenticate: Digest realm="home1.fr",
nonce=A34Cm+Fva37UYWpGNB34JP, algorithm=AKAv1-MD5,
ik="0123456789abcdeedcba9876543210",
ck="9876543210abcdeedcba0123456789"
```

4.4.4 UE Response to the Challenge

The UE will receive the 401 Unauthorized challenge from the S-MIC contained a number of parameters. The received parameters as well as the shared secret allow the UE to generate the response and handle them over to the Gm interface to the P-MICF as a second register message. The UE now needs to add the authorization header to the second REGISTER request including (among others) the following field described in the IMS by Miikka and Georg, (2009).

- The username field – which includes private user identity.
- The nonce field – which is returned with the same value as it was received in the WWW-Authenticate header of the 401 (Unauthorized) challenge.
- The response field – which includes the authentication challenge response that was derived by the UE from the received RAND and the shared secret.

The UE now able to calculate an authentication respond for the authentication request that sent by the S-MICF and generate a second register message toward the S-MICF. As mentioned earlier that the second register request will include the same registration-related information and will be routed in exactly the same way as the initial register request over the Gm interface toward the P-MICF. The P-MICF then forwards the response challenge to the S-MICF. The below information shows the second registration header fields as shown in RFC3310 by (A. Niem.,et al., 2002)

```
REGISTER sip:home1.fr SIP/2.0
Authorization: Digest username="user1 private@home1.fr",
realm="home1.fr",
nonce=A34Cm+Fva37UYWpGNB34JP, algorithm=AKAv1-MD5,
uri="sip:home1.fr",
response="6629fae49393a05397450978507c4ef1"
```

4.4.5 Successful authentication.

As stated before, that all subscribers will have their subscriber profile located in the MSDB. It contains information associated to the users. The S-MICF downloads the user profile over the Cx interface from the MSDB. When the user sends the second register request to the P-MICF, the P-MICF then will check the user request and insure that been sent by the actual user then adds the ‘integrity-protected’ field with the value ‘yes’ to the authorization header and sends the register request toward the S-MICF:

```
REGISTER sip:home1.fr SIP/2.0
Authorization: Digest username="user1 private@home1.fr",
realm="home1.fr",
nonce=A34Cm+Fva37UYWpGNB34JP, algorithm=AKAv1-MD5,
uri="sip:home1.fr",
response="6629fae49393a05397450978507c4ef1",
integrity-protected="yes"
```

The P-MICF will forward the second register request to the S-MICF including the mentioned headers, UE then authenticated by the S-MICF with data provided by the MSDB that was requested by the S-MICF via a diameter Server Assignment Request SAR. The MSDB retrieves the user profile and includes the AVs in the diameter Server Assignment Answer SAA. Once the S-MICF received the SAA it will, match the user request with the AVs, and compares the received response from the UE with the AV that was received in the SAA from the MSDB to check if they match, then it can decide whether the user is allowed to continue with the request or not. If the parameters are identical, then the S-MICF has successfully authenticated the user. The S-MICF now able to generate a 200 OK message to the UE, and the authentication process is completed.

4.5 Session Establishment

This section provides details on establishing a basic session using INVITE messages. The main goal of this section is to provide a description of the process along with a functional requirement for the P-MICF-S-MICF and MSDB when handling the session. This part will also demonstrate the process by providing a call flow and the message description formats for both request and response.

4.5.1 Overview

When the user desires to initiate a session, the UE will formulate an INVITE request. This request will pass through the first point, the proxy, for identification purposes. Then it goes to the server to establish the session and to process the request. Once the server receives the request and downloads the user profile, the server can accept the invitation and send a 200 OK to the UE via the proxy or reject it and send a 300 to 600 response depending on the reason for

the rejection. This general view of the procedure of session initiation has a number of steps to process the INVITE request shown below.

4.5.2 Generating the Request

The user will be able to initiate the session after the registration procedure mentioned in section 4.2. The establishment process starts by sending a SIP INVITE request from the user side. As usual, the P-MICF is the first point of contact for the SIP INVITE message forwarding. When the P-MICF receives the SIP INVITE over the Gm interface, this message contains the URI that is associated to the user and similar to the email address (sip:userA@example.com). This message carries a number of header fields that indicate what methods can be invoked, along with the session description protocol (SDP) in the body of the message that describes the session and includes a set of media and codec that the UE wishes to send. Once the P-MICF receives the invite message, it will forward the SIP message to the S-MICF via the Mw interface.

4.5.3 Processing the Request

After this request has reached the server it will contact the MSDB and send a UAR over the Cx interface for authorization purposes, the MSDB will send back the UAA to the server. The purpose of this step is to download the user profile and authorizes the user. Once that done, the S-MICF will invoke the AS via the Isc interface.

4.5.4 Processing the Response

When the AS received the request, it will then send the result back to the server in order to forward the answer to the user. This stage will pass three main stages to send it to the end user, which are filtering, classification and adaptation. The call flow for the INVITE message is defined in figure 4.9, followed by the SIP header and SDP field shown in figures 4.10 & 4.11.

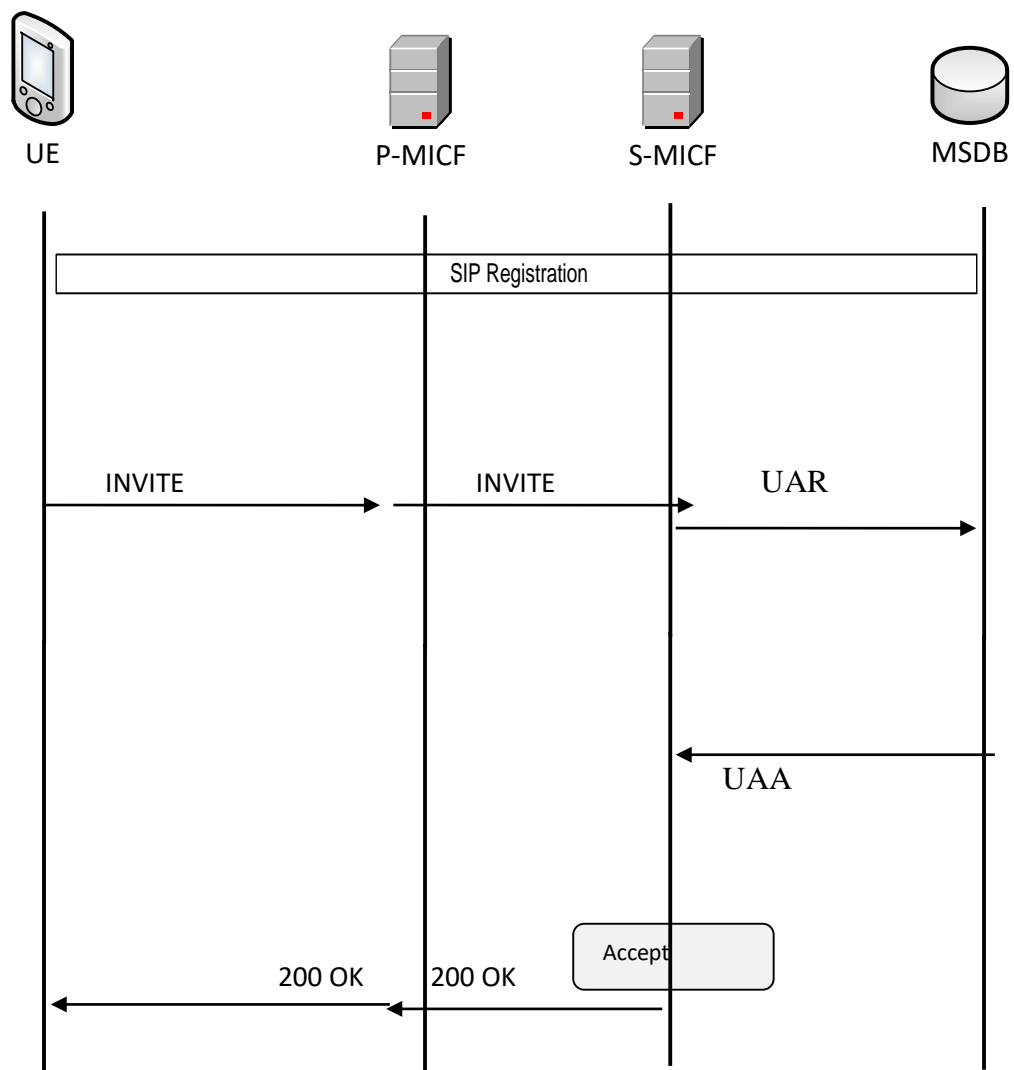


Figure 4. 9 Call flow for the INVITE Message

```

INVITE sip: sip:Alice.Smith@domain.com SIP/2.0
Via: SIP/2.0/UDP p1.domain.com:5060;branch=z9hG4bK543fg
Via: SIP/2.0/UDP ws1.domain2.com:5060;branch=z9hG4bK74gh5
;received=192.0.100.2
Max-Forwards: 69
From: Bob <sip:Bob.Brown@domain2.com>;tag=9hx34576sl
To: Alice <sip:Alice.Smith@domain.com>
Call-ID: 6328776298220188511 @192.0.100.2
Cseq: 1 INVITE
Contact: <sip:bob@192.0.100.2>
Content-Type: application/sdp
Content-Length: 151
v=0
o=bob 2890844526 2890844526 IN IP4 ws1.domain2.com
s=-
c=IN IP4 192.0.100.2
t=0 0
m=audio 20000 RTP/AVP 0
a=rtpmap:0 PCMU/8000

```

Figure 4. 10 Header Field for the INVITE Message

```

SIP/2.0 200 OK
Via: SIP/2.0/UDP p1.domain.com:5060;branch=z9hG4bK543fg
;received=192.1.0.1
Via: SIP/2.0/UDP ws1.domain2.com:5060;branch=z9hG4bK74gh5
;received=192.0.100.2
From: Bob <sip:Bob.Brown@domain2.com>;tag=9hx34576sl
To: Alice <sip:Alice.Smith@domain.com>;tag=1df345fkj
Call-ID: 6328776298220188511 @192.0.100.2
Cseq: 1 INVITE
Contact: <sip:alice@192.0.0.1>
Content-Type: application/sdp
Content-Length: 151
v=0
o=alice 2890844545 2890844545 IN IP4 192.0.0.1
s=-
c=IN IP4 192.0.0.1
t=0 0
m=audio 30000 RTP/AVP 0
a=rtpmap:0 PCMU/8000

```

Figure 4. 11 SDP for 200 OK

The header field shown in both figures 4.10 & 4.11 are from INVITE SIP message, to give the reader a clear understanding of the contents of the message.

4.6 Example Scenario

Now that the procedure of the registration message and the session set up (INVITE) SIP message have been clarified, this section presents two different scenarios for the request data plan and clarifies how the system handles the request over both scenarios. The first scenario is called general search; a non-expert user performs this search, and the second scenario is called the expert search that is performed by an expert user who has done a number of searches and has a record history of the past searches. A detailed description of the request data plan will be demonstrated via data flow diagram for the general user scenario, in addition to the request data flow plan for the expert user scenario, which is supported by the smart layer. This section will look at the way that the system handles the data and how the system invokes the smart layer to adapt the user' request and filter the results in order to personalize the medical search results and offer user preferences data efficiently using user profile. Both scenarios based on the assumption that the users are medical students one of them is registered with the system and with no subscription yet and the second one is an expert user with a record of subscription. It assumes that both users are accessing the system via mobile phone searching for same keywords (**Paracetamol**).

4.6.1 First Scenario -General Search

This scenario is performed by general user who is registered for first time with the system. It is assumed that the user has already registered his SIP URI [sip:userB @example.com](mailto:sip:userB@example.com), and is now able to initiate the session.

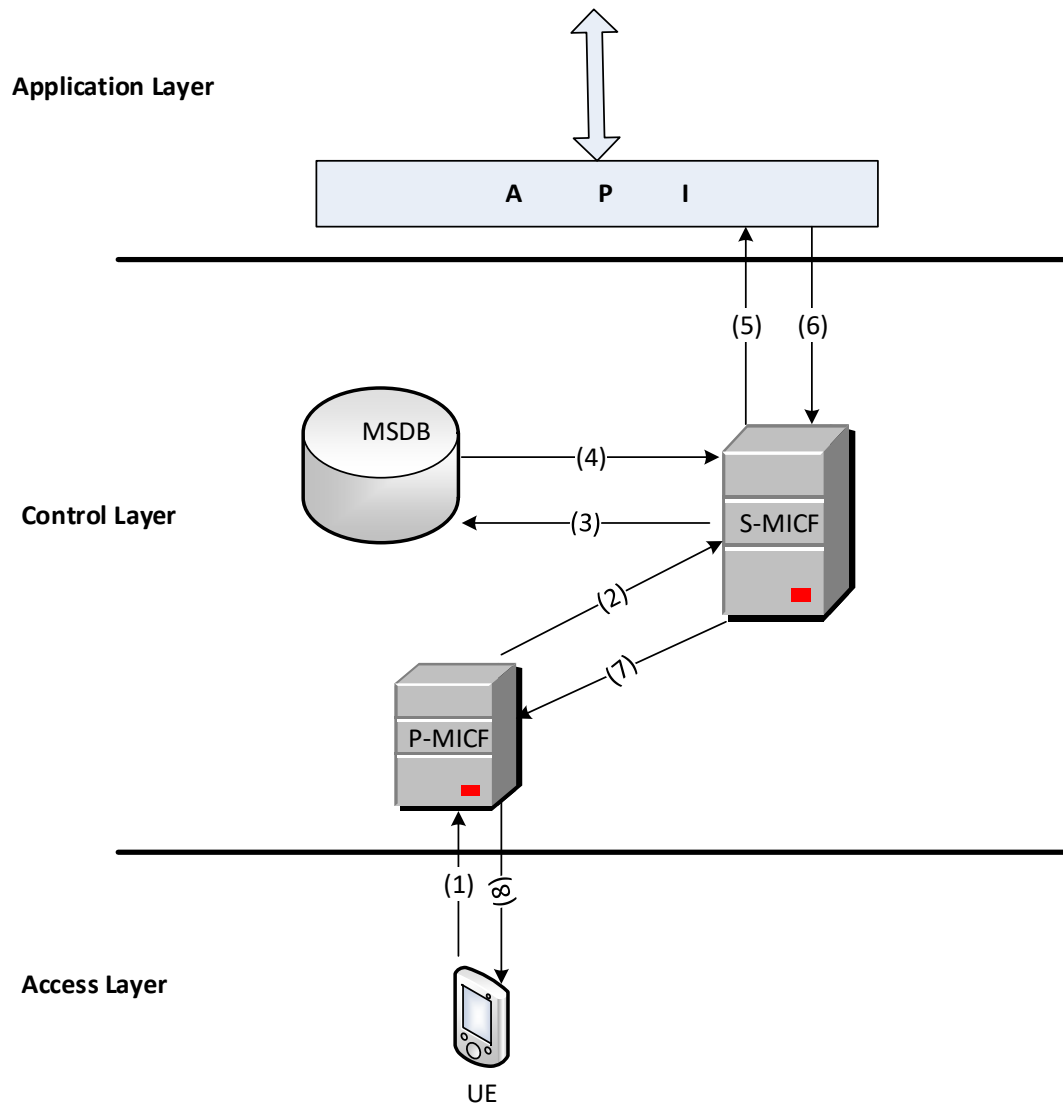


Figure 4. 12 Request Data Flow Diagram Scenario 1

The procedure of the general search is similar to the way that the system handles the INVITE SIP message that mentioned in section 4.3.1. Figure 4.12 shows a call flow for the scenario, which involved the following steps:

- 1) The user will send a SIP message towards the P-MICF over the Gm interface with session description protocol SD.
- 2) Then the P-MICF sends the INVITE request towards the S-MICF over the Mw interface.
- 3) The server will contact the MSDB and send a diameter Location Information Request (LIR) towards the MSDB via the Cx interface in order to download the user- related information.
- 4) The MSDB now will retrieve the user information and download the user profile that contains the user location information, the MSDB then sends a diameter Location Information Answer (LIA) in the same way via the Cx interface to the S-MICF.
- 5) Based on the information retrieved from the user profile, the server is now able to recognise who is searching (type of user). The server will learn from the information stored in the user profile a number of pieces of information, for example: the user subscription, the number of searches done by the user. Now the server has identified the user, after performing services-related checks based on the subscriber profile, therefore it will forward the INVITE request to the AS via the Isc interface.
- 6) Once the SIP AS generates the 200 OK it will send it to the S-MICF.
- 7) The S-MICF then handles the 200 OK response back to the P-MICF via the Cx interface.
- 8) Finally, the P-MICF will send the response back to the UE over the Gm interface.

4.6.2 Second Scenario Expert Search

As in the first scenario, the user has to register his identity sip:userA@example.com and then send the request message to the system.

The data plan of this scenario for the expert user will follow the same process of the first scenario that was mentioned before but in this type of search the smart layer functions will be invoked to support the user's request in order to adapt and customize the user's information needs. Figure 4.13 shows the data plan process for the scenario, which involved the following steps:

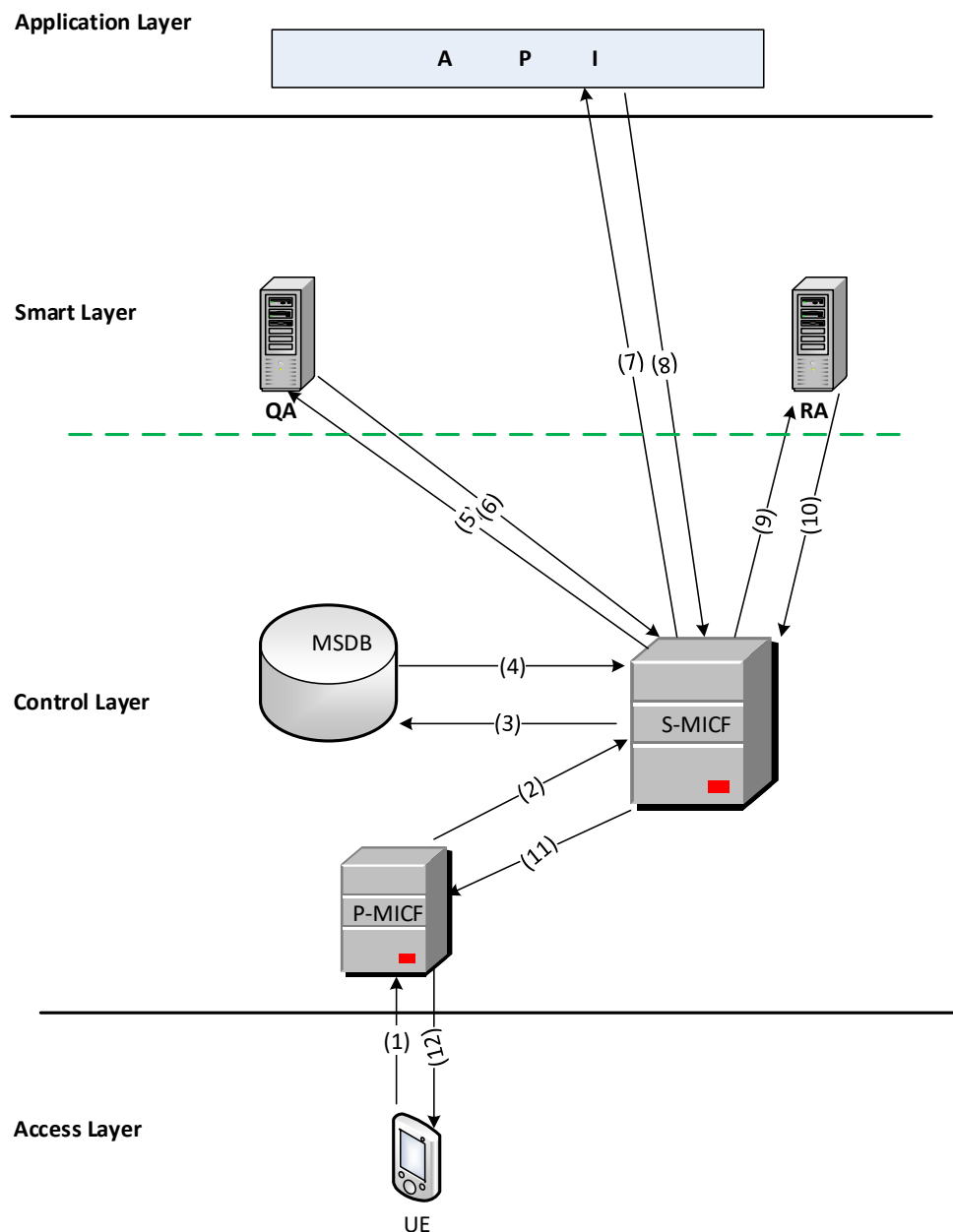


Figure 4. 13 Request Data Flow Diagram

- 1) The user will be able to send the request via a Gm interface to the main point in the system
- 2) Once the request is delivered to the proxy, P-MICF it will forward the request to the server S-MICF through the Mw interface.
- 3) When the S-MICF receives the request, it has to contact the database MSDB in order to authenticate the user identity and user location information. Therefore, it will send a diameter Location Information Request (LIR) towards the MSDB via Cx interface.
- 4) The MSDB now will retrieve the user information and download the user profile that contains the user location information; the MSDB then sends a diameter Location Information Answer (LIA) via the Cx interface to the S-MICF.
- 5) Now upon information received from the MSDB, the server will identify the user type, it will then contact the smart layer (Query adaptation) over Mw interface to support the request, and this can be via a number of preferences, which already described in chapter 3.
- 6) Once the smart layer adapted the request, it will forward it back to the server through Mw interface.
- 7) The server then be able to forward the request to the Application Programming Interface API over Isc interface.
- 8) The web search results will be back from the API to the server over Isc interface.
- 9) When the server receives the results, it will contact the smart layer for the second time to apply filtering and customise the search results, therefore, it will invoke the Search adaptation unit over Mw interface in order to adapt the results based on the user's profile this step is described in chapter 3.
- 10) Once The Result adoption has filtered the result based on the user profile, it will send the results back to the server through Mw interface.

- 11) The server then will forward the result to the proxy via Mw Interface.
- 12) Finally, the proxy will send the results back to the user via Gm interface.

4.7 Summary

In this chapter, a functional evaluation of the signalling flow plan required some explanation of the contents of the message. Therefore, a description of the SIP message contents are defined. The two types of the SIP message methods name and meaning that related to the request and response are illustrated along with the header fields that both message (request and respond) contain are included and an example of message header filed is given to clarify the filed in both massage. In addition to the session description protocol that carried out by the SIP message body which provide further information to describe the session is illustrated. This chapter also demonstrates the registration process along with the authentication procedure. Moreover, an overview of the session establishment and handling the request is described. Finally, examples scenarios of two different user (general user and expert user) are given, both scenarios are demonstrated with data flow diagrams.

Chapter 5

Smart Layer Design Methodology

5.1 Introduction

The smart layer has mentioned earlier in chapter 3 created based on a number of steps that confirmed the needs of the smart layer to be introduced on the proposed framework, this chapter will introduce the smart layer design methodology and the steps that led to the need of the smart layer to be introduced in our framework. Figure 5.1 shows the organizational diagram of the smart layer design methodology structure.

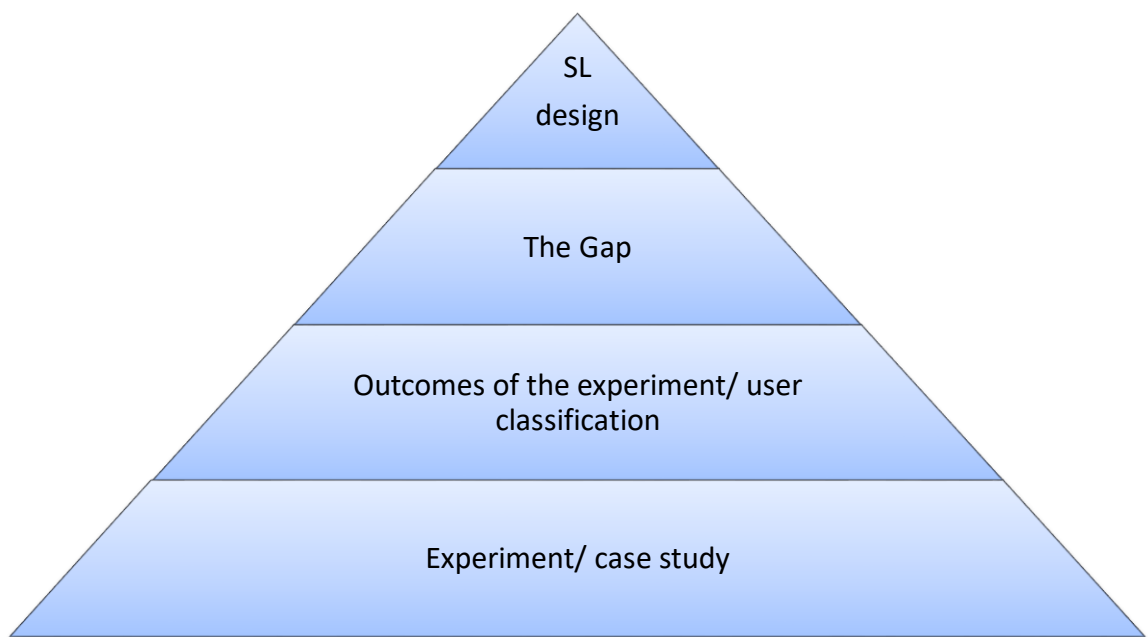


Figure 5.1 Organisational Diagram of the SL Design

The structure represents the logical flow of the followed method to designed the Smart Layer which includes the following steps:

- 1- A case study applied to identify a number of the results obtained by different type of users from medical domain under three subjects.
- 2- User classification has been identified from the experiment (General user, Expert user).
- 3- The experiment has identified the gap of the need of filtering system that filtered out the irrelevant results based on the user profile.
- 4- The smart layer introduced based on the above steps.

5.2 Case Study

This part presents a case study based on the number of results retrieved by a number of users from a web search engine and medical databases in response to their queries. The study tested the validity of the assumption on which the framework was designed and created, and those results will be presented later in this chapter. An experiment was carried out by different types of users within medical community from the Heath hospital in Cardiff searching for three keywords given by the researcher on their preferred search method. The experiment shows that all the users performed their search using the following search options: the generic search engine Google, Medline and the PubMed (both of which are medical databases listing professional biomedical literature) and, finally, MedlinePlus, which is the National Library of Medicine's web site for consumer health information. The following search methods were also used in this experiment: Wiley, ScienceDirect and the ACM digital library, but those results were not included in this experiment because the number of participants who carried out

searches were too low (only 5 participants; 2 nurses and 3 medical students) and the results obtained by them were low also. Consequently, the low number of results for these three databases in comparison to the other search engines means that the findings were statistically irrelevant so they were excluded. Figure 5.2 below shows the search results obtained by both of medical students and nurses under the subject of diabetes, which was the only subject that both users used the three excluded databases Wiley, ScienceDirect, ACM along with the other databases that other users have performed their search on them (Google, Medline, MedlinePlus and PubMed).

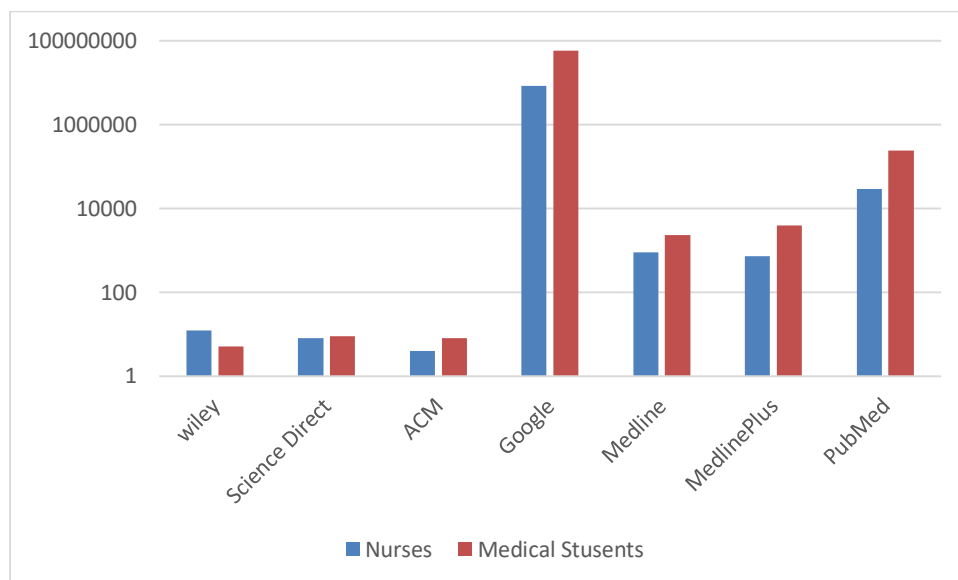


Figure 5. 2 Search Results Obtained by Nurses and Medical Students for Diabetes

The results of the experiment will be presented in two methods: first is a table and second is a figure. In the figure, the X (horizontal) axis shows the search engine method used, and the Y (vertical) axis shows the exact number of results. The graphic representations will follow the same structure for all groups of participants.

5.3 Design and Setup of the Experiment

5.3.1 The Purpose of the Experiment

This experiment was conducted to investigate the way that physicians search their information and the number of results they obtained from different preferred methods under a number of medical search keywords given by the researcher. The outlined step is done in order to identify a way to customize and reduce the amount of results returned and to provide a filtering method to avoid a large number of irrelevant information. This experiment will outline why is the smart layer designed. The design of the experiment is outlined in the subsequent subsections.

5.3.1.1 Selection of querirs

The first step was the query selection. Three keywords were selected as the top three results returned by the different types of users from the healthcare community in responding to the question (What are the most five interesting keywords you do usually use to search). These keywords were: Diabetes, Osteoporosis, and Blood pressure. The question submitted online using Survey Monkey and the results returned online within three weeks. 43 participants were participated in this part of the survey the majority of them were medical students with 20 participants and 7 nurses, 5 pharmacists, 4 midwives, 3 pediatrician , and 4 dermatologists.

Figur 5.3 below shows that diabetes was the most popular topic that 41 respondents agreed to search for followed by 38 respondents who search for blood pressure and finally 30 respondents who have chosen the osteoporosis. The participants were also interested in other subjects, such as cancer, thyroid and heart disease. Those topics did not select to be the main search queries of this experiment because of that the number of the respondents was low compared with the three topics mentioned before.

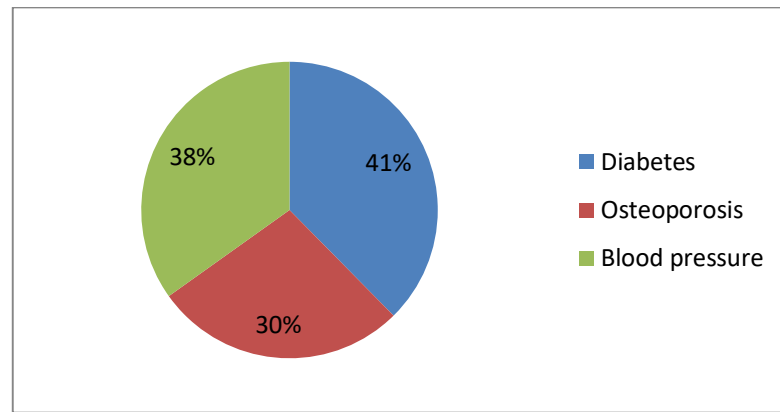


Figure 5.3 The Three Search Queries

5.3.1.2 Web search

In order to do the web search, the participants were contacted via individual emails and those accepted were considered for the questionnaire. Seventy-three users were requested to search under the three selected queries given on their preferred methods and to write the number of results received from each one. 43 of physicians did actually take part in the survey out of which 13 Nurses 11 Medical Students 9 Pharmacists and 10 Dentists. It is important to note that the search methods used in this experiment are the most common methods that all users agreed to use. However, there are other methods but did not show in this experiment because the number of the users was so low, almost 5 of the participants used them. These methods are Wiley, ScienceDirect and ACM digital library. The first part of the experiment was done by the researcher, the second part was done by all 43 of the participants and, lastly, an additional search was done by 4 dentists who were selected by the researcher in order to perform the expert serch.

5.3.1.3 Expert search

Based on the analysed results obtained from the user's group, the researcher selected a small number of participants to perform more specific search on one subject selected by the participants themselves. They were then asked to target more specifically that subject in order to obtain a lower number of results. Four dentists were willing to participate in this part of the experiment. They were asked to choose a keyword of their interest and to perform the search using their preferred search methods. They were informed of the purpose of this step, which aimed to reduce the number of results by selecting suitable keywords in order to obtain fewer results . This step conducted to show what the smart layer is designed for. The four dentists worked together to perform the search, which focused on oral diseases, and then they tried to narrow down the search and obtain in-depth details for specific information about this subject.

5.3.1.4 Data Collection

As indicated earlier, the questionnaire method was adopted for the research. It was presented to the participants via email with a set of search questions to perform and provide the pattern of a number of the results obtained for each search platform. Using the responses, the data analysis was conducted. The results of the four group of users returned to the researcher within six weeks. And the second part of the survey (the expert search) was returned 2 weeks later. There were four search methods that most participants agreed to perform the search on them that are: Google, Medline, Medline Plus and PubMed.

5.4 Results of the Experiment

5.4.1 General Search

Initially, the researcher has conducted the search using the same search methods that all the participants also used. The main purpose of this step was to provide the researcher with the possible outcomes of the search that might be conducted by a general user who could be a patient, a student or anyone interested in searching for medical information. The researcher used the exact keywords and wrote the number of results that appeared at the top of the first page returned by the search methods, as shown on figure 5. 4. And this is the way that all participants followed to record the search results .

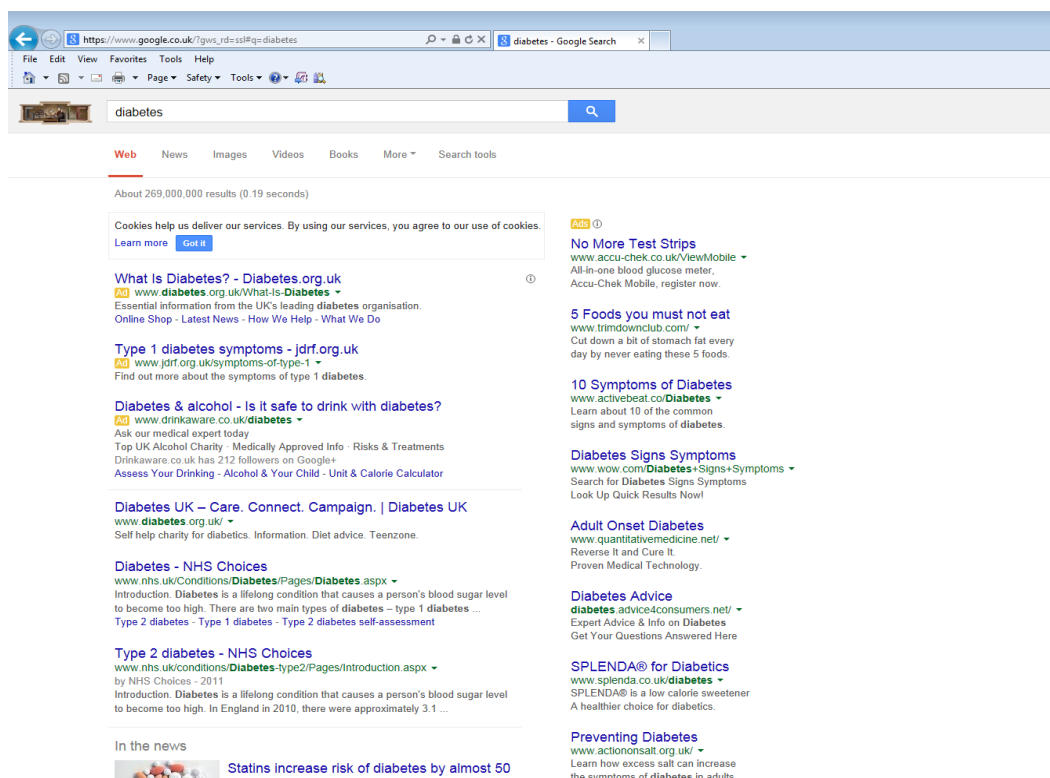


Figure5. 3 General Search Results on Google

However, some of the search results might be irrelevant to the main topic such as the information regarding services or products that might be offered by the search engine based on the search topic.

The number of the search results for all three topics is shown in table 5.1.

Table 5. 1 Number of General Search Results by Search Method

Keywords	Google	Medline	MedlinePlus	PubMed
Diabetes	269,000,000	451,498	5,670	489,522
Blood pressure	11,700,000	375,208	6,014	508,446
Osteoporosis	3,260,000	65,056	1,028	65,116

As indicated in table 5.1, Google had the highest number of results under Diabetes at 269,000,000, followed by PubMed at 489,522, while MedlinePlus had the lowest number of results at 5670, which is still quite high. The researcher did not use other terms that can be listed under the name of search topics as results, the returned results were very high and that shows an indication of the effect of selected terminologies on the search results. Figure 5.5 shows the obtained results from Google to make that comparison clearer for the reader.

The outcome of this search shows the following indication:

- Google may provide fewer results for the participants by using different terminology and the advanced search.
- Participants may use other terms to perform the search in order to get less results than the researcher had.

- Medline Plus may retrieve fewer results than the other since it is a website and has a limit linked to other databases.
- Because of that both of Medline and MedlinePlus are databases which provide search features such as (AND ,OR) the participants will benefit of these search preferences.
- This search indicates to the ease of use for all of the search methods used to perform that search using simple queries which mean patients and ordinary user are able to search them.
- The use of more simple queries appears to have a low impact on the results retrieved.
- Due to the high number of results increasing effort on learning how to perform the search using advanced search and other feature is important for the users to get the right and relevant information.

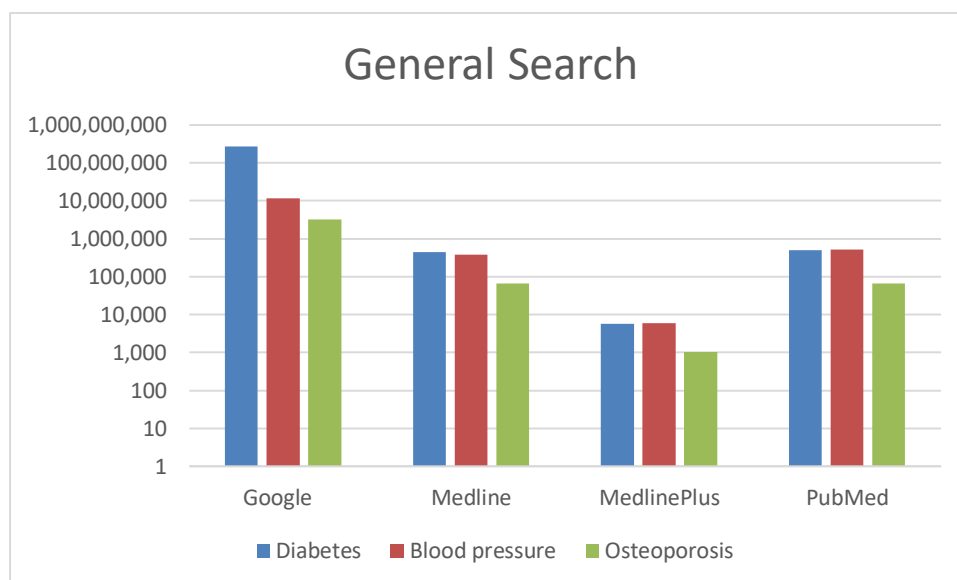


Figure5. 4 Results Obtained by the General User

In the second part of the experiment which performs by four groups, the participants were asked to conduct a web search to investigate some information under each subject and to write the number of returned results from each database and search engine that they used. Although all the participants used the same keywords, they searched for different aspects under each subject and they received different results based on their interests. The following section will discuss the search that was conducted by four different groups of nurses, dentists, pharmacists and medical students.

Before providing the results, it is important to clarify the main purpose of this step, which was to identify a way to reduce the number of results returned by a search method based on practical user experience. Therefore, at this stage the participants were asked to be more specific in selecting their information in order to reduce the number of results and to avoid a large amount of irrelevant information. As mentioned before, all the participants searched almost all of the same databases:

Google, Medline, MedlinePlus and PubMed

In order to provide as much information as possible, the results of the experiment have been classified based on four groups of participants.

5.4.2 Group 1 (Nurses)

This is the first group of this experiment contains of 13 nurses who performed the search. Table 5. 2 illustrates the search results obtained by the nurses 'group searching all three selected keywords.

Table 5 . 2 Search Results for the Nurses

Subject	Keyword Search	Google	Medline	MedlinePlus	PubMed
Diabetes	Diabetes diagnosis and management	8,240,000	893	723	28,636
Blood pressure	Blood pressure management in primary care	8,230,000	19	340	2,919
Osteoporosis	Osteoporosis assessment and monitoring	2,160,000	49	35	292

Table 5.2 shows that although the nurses have a clear orientation toward medical databases, especially Medline and PubMed, the obtained results show that participants have utilized Google to search for their information although that Google returned a larger number of results in comparison to the other search methods, which makes it the most appropriate method for this group of participants. It is typical that Google retrieves a huge information related to each subject because Google is an open source search engine that provides access to full-text material as well as multimedia whether this information is related to medical information or products advertisement or the latest news about the topic being queried. In contrast, the Medline and PubMed databases provide full-text access and in some cases, provide an abstracts only and sometimes required the user to subscribe to the service and pay for full access, which make them a second choice option for this group.

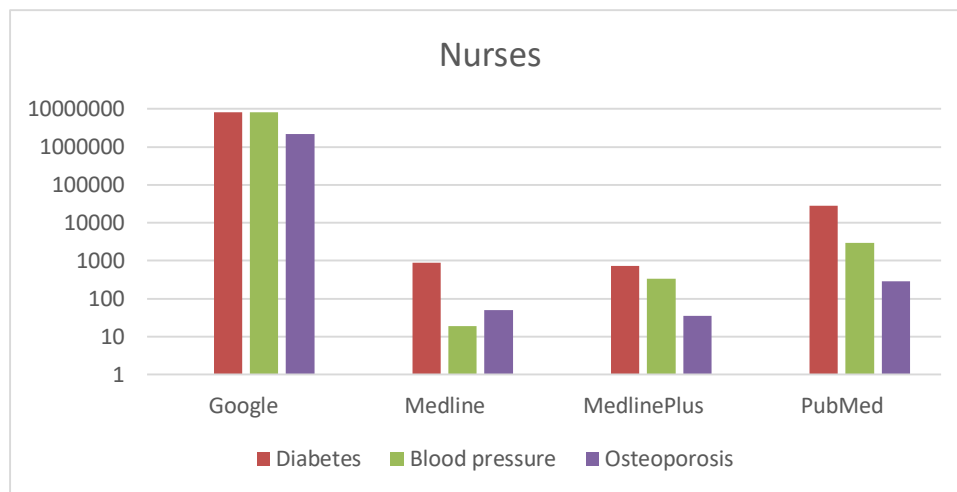


Figure 5. 5 Search Results Obtained by Nurses

A comparison showing in figure 5.6 between the Medline and PubMed results for diabetes shows an evidence that the type of search methods had a significant impact on the results even though both Medline and PubMed are considered to be two of the most reliable databases in the medical domain that provide articles from academic journals, access to online books and information about biomedicine and health statistics. The decreases of the obtained results for the diabetes results obtained from Medline compared with PubMed results shows evidence of the role of the search operator that offered by the Medline. However, the results confirmed that the keywords used had a low impact on the number of results returned.

5.4.3 Group 2 Dentists

The second group of this experiment contains of 10 dentists and have performed the same search as the first group. Although all the participants in the group of dentists were aware that Google is an open source search engine that provides access to a large number of documents and retrieves a large number of results, the information presented in table 5.3 illustrates that Google was the search engine option that was most frequently used by the participants in this

group and it was also the search method that delivered the most results for each of the searched topics.

Table 5. 3 Search Results for the Dentists

Subject	Keyword	Google	Medline	MedlinePlus	PubMed
Diabetes	Diabetes and oral health problems	174,000	142	41	206
Blood pressure	Managing hypertension for oral health	4,800,000	22,766	86	24
Osteoporosis	Osteoporosis and tooth loss	144,000	177	61	216

The analysis results presented in table 5.3 shows that the searches done on Google had the largest number of results of all of the given keywords: 4,800,000 results for managing hypertension for oral health and 174,000 results for oral health problems. In contrast, MedlinePlus only gave 41 results for diabetes and oral health problems. A comparison shows in figure 5.7 of the results from both Medline and PubMed shows a slightly significant difference for osteoporosis: 177 results on Medline and 216 results on PubMed; still, these results were generally much lower than the results obtained from Google. With the exception of managing hypertension for oral health, MedlinePlus shows the least number of results for the subjects investigated.

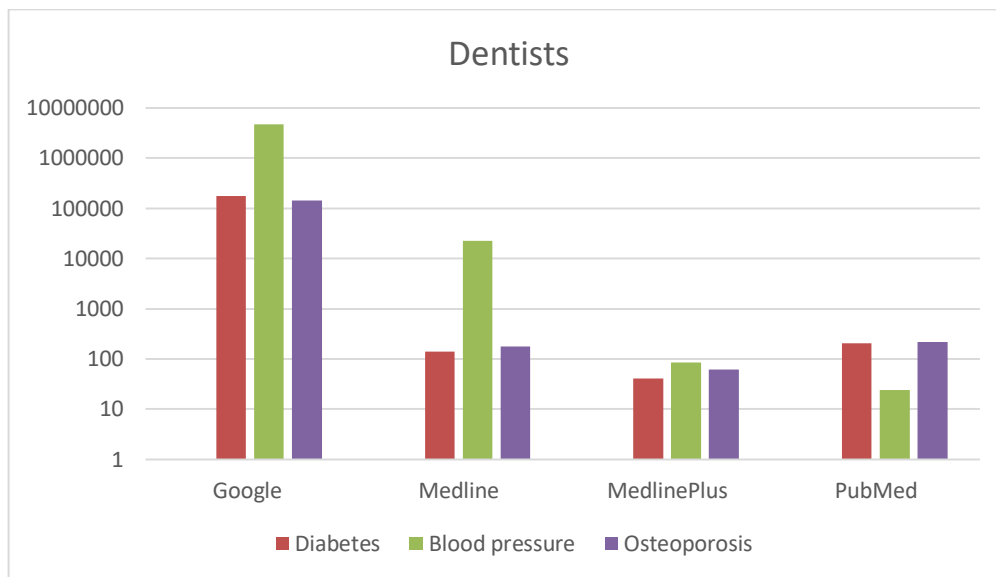


Figure5. 6 Search Results Obtained by Dentists

Because of that Google was most popular methods in this group ,the impact of the search keywords were low although that other methods returned fewer results than Google but they did not give a clear indication of the impact of the keywords on the results due to the convergence of the results on them on all three keywords search which were almost under 1000, except the blood pressure which was higher than 10,000 on the Medline.

5.4.4 Group 3 Pharmacists

This group is the third group who participated for the purpose of this experiment and contains 9 pharmacists who search for the three keywords given.

Table 5.4 shows that pharmacists more likely to search for drugs, alternative therapies and medical devices than other groups, and because of that Google may provide significant information related to the drugs and medical devices it was one of the convenient sources that pharmacists used to search for the three topics investigated.

Table 5. 4 Search Results for the Pharmacists

Subject	Keyword	Google	Medline	MedlinePlus	PubMed
Diabetes	Monitoring diabetes	39,200,000	1,647	644	16,861
Blood pressure	Blood pressure device	18,700,000	951	1,185	9,579
Osteoporosis	Vitamin D ₃	13,200,000	10,327	38	26,641

In this group, the evidence clearly shows that Google had the largest number of search results for all of the search topics, and by taking Google out of comparison (due to the huge results returned). A comparison shown in figure 5.8 between the two databases (Medline, MedlinePlus) shows that pharmacists retrieved fewer results from Medline than the MedlinePlus as an example of that the osteoporosis PubMed retrieved around (26,641) results which are almost double the number of Medline results (10,327), the difference was significant. Although that MedlinePlus returned the least number of results for all of the keywords in comparison to the other search methods but that can not clear evidence of the impact of the keyword on the search results because MedlinePlus is a medical website. Generally speaking in this group there was a slight impact of the query on the search results, and the Medline database might be the most appropriate search methods that can provide less information compared to both of Google and PubMed results.

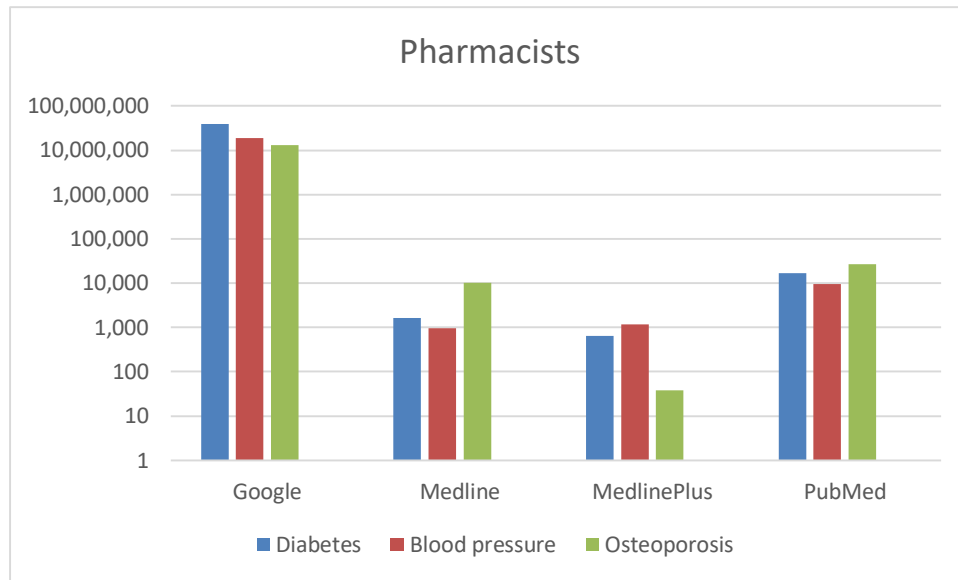


Figure 5. 7 Search Results Obtained by Pharmacists

5.4.5 Group 4 Medical Student

The fourth group of this experiment is performed by 11 participants of medical students. As the aforementioned groups this group has also performed their search using the same search queries and obtained the following number of results shown in table 5.5.

It is clearly that medical students use Google as the first option to search for their information because of that Google is a gateway that supports students with information related to learning and finding medical terminologies, on the other hand it can be a huge effort for the student to browse the relevant information from Google and that can be proven by the results they obtained from Google for all three topics investigated in this experiment which shows that Google returned the largest amount of results for all topics over all the groups participated in this experiment.

Table 5. 5 Search Results for the Medical Students

Subject	Keyword	Google	Medline	MedlinePlus	PubMed
Diabetes	Diabetes symptoms	58,000,000	2,279	3,891	237,495
Blood pressure	Blood pressure control	96,900,000	16,961	3,404	119,891
Osteoporosis	Osteoporosis screening	6,340,000	947	184	24,495

Across the board, Google returned a larger number of results than the others. For all three keywords, it can be clearly seen from figure 5.9 below that the Medline results are much lower than the PubMed returns in every category (which are large, but significantly smaller than the Google returns). However, the students may get the benefit of the search feature offered by the Medline. Although of that Medline database retrieve fewer results than the PubMed. The experiment shows that there was a slight impact of the keyword selected on the search results and that might be related to the lack of experience on performing medical-related information or selecting medical terminologies related to the subjects investigated.

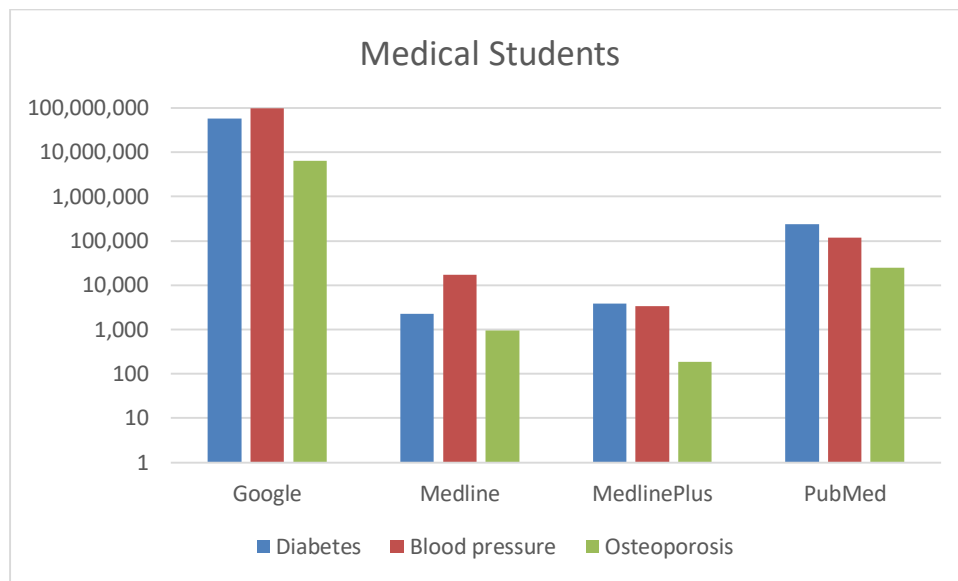


Figure 5. 8 Search Results Obtained by Medical Students

5.5 Discussion

Although all the study participants performed their queries using the same search methods, the number of results they received was different depending on the keywords they used. Because this investigation has looked at the impact of using different types of keywords has on the number of results , this study's findings have identified several correlations between the search results obtained by the four study groups.

In general, there was a quantitative similarity in the number of results for the four groups under all the three subjects searched using Google; the Google results numbered in the millions compared to the results of the other search methods. In contrast, there was no major similarity among the obtained PubMed, Medline results for all four group. Across all of the groups participated PubMed considered to be the second source that retrieved a lot of results after Google while MedlinePluse was in the first place which returns less amount of results for nurses, Dentists and Pharmacist. Furthermore, for all groups participated the diabetes, blood

pressure and Osteoporosis confirmed the impact that using different search keywords on the the number of results and that can be seen from the results which shows that blood pressure has recorded the highest number of results for the medical student and dentists while osteoporosis has recorded the least number of results for nurses and medical students. In contrast, it Osteoporosis has retrieved the highest search results for the Pharmacists group.

By comparing the two databases Medline and MedlinePlus the analysed data confirmed that the number of results obtained from both were not similar on all search topics. Although both are medical databases designed to retrieve medical information but the experiment shows different outcomes for each group, an example of that the nurse's search under diabetes PubMed returned around 28,636 which is the trouble number of results that Medline retrieve. On the contrary, PubMed returned the less results under blood pressure topic which performed by the dentist group comparing with Medline results. Moreover, the number of results under diabetes and osteoporosis for the nurse and pharmacists groups were nearly close to each other on the Medline and MedlinePlus, which means that either the keywords were too similar or the participants (nurses and pharmacists) or may have had the same area of interest.

The Medline and PubMed and MedlinePlus results show clear evidence for how the results are affected by different search words in all three topics/subjects. Given that results were obtained for all the participants, it can be clearly seen that the keywords that were used to perform the search had a significant impact on the results. This means that the keywords determined if the results were related to the query or if they fell far short of the ideal. Thus, it is important to consider the user's preferences and that is the purpose of designing a smart layer for search engines.

As consequence to the information obtained from the previous searches done by the participants in the four study groups and based on the assumption made in designing the framework, which aims to support the search and customize the results in order to get less and related results that achieve user satisfaction. The four study groups confirmed that how different users performed the same search queries and retrieve different number of results based on the keywords selection, it is also confirmed that the results obtained from all search methods still high. It is important now to investigate the way in which the participants use their preferences to get less results, doing so will help increase the likelihood of obtaining relevant results and narrowing down the total number of results delivered. Therefore, the next part of this experiment seeks to prove the assumption that the number of results will meet the user's needs by increasing the number of searches using different keywords and preferences. This part of the experiment is called the expert search, because it is done by users who have more experience conducting this type of search.

5.6 Expert Search

Based on the analysed results obtained from the first part of the experiment, it is clear that most of returned search results are still large, even though specific databases in a medical domain were used. Therefore, the researcher selected a smaller number of participants and asked them to perform a more specific search of one subject selected by the participants themselves. They were then asked to more specifically target that subject in order to obtain a lower number of results. Four dentists were willing to participate in this part of the experiment. They were asked to choose a keyword that interested them and to perform the search using their preferred search methods. They were informed of the purpose of this step, which aimed to reduce the number of results by selecting suitable keywords in order to obtain fewer results. The four dentists worked together to perform the search, which focused on oral diseases, and

then they tried to narrow down the search and obtain in-depth details for specific information about this subject. Table 5.6 shows the minimum number of results obtained after 20 attempts.

Table 5.6 Expert Search Results

Keywords	Results
Oral diseases	11,200,000
Periodontal disease	10,456,000
Periodontist	7,860,000
Classification of periodontitis	2,105,000
Severe periodontitis	767,000
Risk factors for periodontitis	493,000
Diabetes and periodontitis	571,000
Management of diabetic patient in dental clinic	252,700
Extraction of hopeless teeth	194,400
Post extraction management	92,160
Post extraction bleeding management	7,841
Management options of post extraction bleeding	3,710
Management of post extraction bleeding gum	1,300
Management options of post extraction bleeding	798
Type of post extension	540
Post extraction bleeding causes	371
Risk factors for extraction bleeding	676
Unusual case of post extraction bleeding	260
Management of post extension	145
	123

As can be clearly seen from the information presented in table 5.6, there was a significant difference in the number of returned results in comparison to the results from the searches obtained in the first step of this experiment (with the four groups). The information in Table 5.6 shows a gradual decrease in the number of results returned by Google from 1 to 20 sessions.

Surprisingly, the participants agreed to conduct their search using Google and they did not use any medical database. At the beginning, the four dentists selected oral diseases (generic topic) as their query. Consequently, that session delivered millions of results, which gives an indication of how the results are impacted by the terms the user selects. Therefore, in the second session the dentists tried to focus on selecting more suitable keywords that could have a high impact on the results. After the third attempt, the number of results gradually started to decrease gradually. Although the number was still large for the management of post extraction bleeding option, the outcome still shows an improvement over the previous attempts and indicates that it was possible to obtain fewer results after 16 attempts under the post extraction bleeding option.

At session number 20, the number of results was much lower than before, that can be clearly seen in figure 5.10. Although the participants selected more accurate keywords under the oral diseases option, the outcome shows that the lowest number of results was 123 on the type of post extension haemorrhage option, which still a large number.

The above step confirmed that increasing the number of searches by selecting more accurate keywords made the results more powerful, but that did not satisfy the user's needs in this experiment due to the high number of results obtained in last session.

Undoubtedly, accuracy in selecting suitable search queries has a strong influence on search results and that was proven based on the outcomes from the third to the 20th sessions in the

expert search portion of this experiment. However, the experiment proves that there is an essential need for a system that filters and customizes the final outcomes so that fewer results are provided. This proves that it is important to design the framework to be more specific and intelligent so that results can be customized based on each user's preferences.

The expert users also confirmed that response time is of great importance especially when they need the information for decision making regarding the patients' health or for treatment and prescriptions.

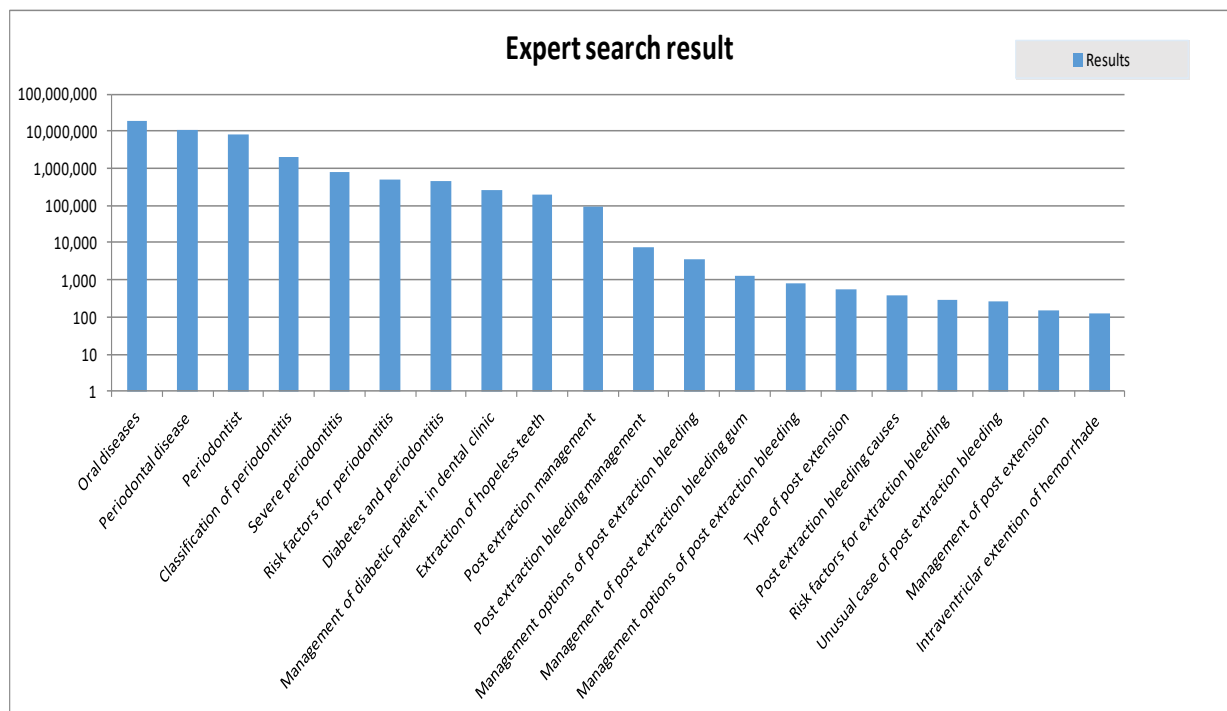


Figure 5.9 Expert Search Results

5.7 Highlights of the experiment outcomes

The experiment indicated that all of the participants used Google for their query search, even though it delivered a larger number of results than Medline, MedlinePlus, and PubMed. The

first part of the evaluation tested the impact that using different types of keywords would have on the number of results, and the findings show a comparison between the returned results for all of the users, which demonstrates a significant difference in the number of results returned by the medical databases (lower) in comparison to Google (higher); no major changes were found between the results for the medical databases. The second part of the experiment investigated the impact that the user's preferences would have on the results.

The findings indicated that the user's experience of performing the search has an impact on the returned results. This is due to the user's ability to specify and narrow the search criteria by modifying the keywords to set the desired query. This also has the potential to reduce browsing time. For this outcome to occur, the user must be familiar with using the specific medical databases, especially since these methods are designed to cover medical information. This means that the search would probably return a list of the most relevant results; however, that was not confirmed in our experiment. In fact, it is contrary to the search results obtained in the first part of the experiment which found that the participants were more familiar with Google than they were with other search methods.

Another assumption was that the users might change the way in which they search for information once they obtained a high number of results. However, the analysed results show that this change in behaviour did not happen with the 43 participants who still utilized Google to obtain information.

A third assumption was that increasing the number of searches using different preferences would decrease the number of results and make them more relevant. This idea not only confirms that the user's preferences are important, it also reinforces the significant conclusion for the need to find a filtering system that can help customize the results in better way, and that is why the smart layer should be designed.

As we assumed, users need a smart system that has the ability to learn from different types of user preferences, which was proven from this experiment. This kind of smart search will provide higher satisfaction because it will provide a lower number of result . In Chapter 4, we described the flow call and the communication between each layer. In this chapter the users were tested to determine whether or not the assumption made about the need to build the framework and the need to design the smart layer is valid, and based on the results from the experiments the users have confirmed that this assumption is correct.

5.8 The Smart Layer Description

The existing search methods whether are a medical database, search engines, and medical websites retrieve the results based on the information supplied by the user as query and not based on the user interests. As consequence, the user sometimes may find some of these results irrelevant to the query or may spend more time to go throw the retrieved pages to get in to the right information. Therefore, filtering and re-ranking the search results are needed in order to avoid the vast amount of returned information. Moreover, the search results should be filtered based on each user's interests which is stored in the user profile along with other information related to the user (e.g demographic information, subscription..etc). For that purpose, this study has introduced the smart layer to customize the search results for different users based on their interests, preferences, and information needs (user profile). This part will give a detailed description of the smart layer and look at the way that the smart layer is being used to personalize the medical search results and offer user preferences data efficiently using user profile.

5.8.1 The Aim of the Smart Layer

The smart layer is the third layer in the framework located in the main connectivity unit MICF this layer considered to be the main feature in the framework which is designed for the purpose of customizing the search results. The SL aims to personalize and adapts the search results to the users based on their profile in order to improve the relevance of the search results at lower number of results and avoid the large number of information it will customize the information for each user according to user- needs. In addition to learning and understanding user's interest from the past search (search history).

5.8.2 The Design of the Smart Layer

Search personalisation is a technique used to filter and re-rank search results to present the right information to a particular user at the right time. And because of the variety of specialism in the medical filed each user has a distinct background based on the specific medical field and different area of interests. When those users seek information a personalised search result is needed to meet the individual needs. Therefore, the smart layer introduced as a solution to customize and tailor the search results to a particular user based on that user's interests and preferences. This layer will play the main role of personalizing the search results and can be achieved via two main steps first, is learning the user profile (user interest, user search history, keywords query and user subscription) and second is customizing the search results (filtering and ranking). The two steps are performed by two main components designed for the already mentioned purposes.

5.8.3 Smart Layer Components

As mentioned before that the SL will learn from the profile of the user and will rank the results based on that, for that purpose the smart layer contains of two main components each one has created to perform different tasks:

5.8.3.1 Query Adaptation Engine (QAE)

The query adaptation engine aims to enhance and support the users' query received from the server S-MICF with related user preferences stored in the user profile for the purpose of providing personalized information. Moreover, this unit is to allow the search query to be constructed, using search terms that would identify methodologically relevant results to be retrieved. This unit is responsible for the following tasks:

- Understand and support the user's query based on the index and the term of the medical knowledge (MeSh).
- Customise the user's query. This unit has the ability to improve the query and modify the query to be more customised to the specific user profile via having more specific information that matches the actual query based on the stored profile information in the MSDB. For example if the user is a medical student who previously done number of search in the area of (tooth decay – children) and later searched using a different keyword (gum bleeding or tooth decay) this unit then match the user profile and will learn from the profile that this user interested in gum and tooth disease in kids only then it will send a suggestion back to the SMICF server to have more related keywords to the query such as (gingiva bleeding diseases and children).

- Learning the user profile and browsing history that will be describe later on in this chapter.

5.8.3.2 Result Adaptation Engine (RAE)

Result adaptation engine is a filtering system that aims to customize the search results in order to get a relevant information and to minimise the number of results to avoid the irrelevant information. Results customization will be based on the user profile and from the number of sessions that requested by the user (search history). Based on the user profile the system will learn the users' needs and rank the results based on that. Ranking the results will be based on a number of preferences that will be described later on this chapter. The RAE unit is responsible for:

- Filtering and ranking the results based on the user profile in order to obtain better accuracy and relevant results. The ranking should be used in conjunction with selected preferences that the user has already set and stored them at the database MSDB unless if that the user has subscribed and paid for the services in this case the top results will be for the paid services followed by the most relevant information that match the user preferences.
- accuracy and relevant results. the ranking should be used in conjunction with selected preferences that the user has already set them and stored at the databases MSDB unless if that the user have subscribed and paid for the services in this case the top results will be for the paid services followed by the most relevant information that match the user preferences.

It worth mentioning that both entities of the smart layer, QAE and RAE, are separated here for demonstration purposes to define the functionality for each unit. However, in the evaluation chapter both entities will be merged as part of one single cache server for

evaluation purposes, in which the server will be handling both the requests and responses into and out of the smart layer.

5.8.4 Smart Layer Search Parameter

For the purpose of personalization, the smart layer gives the user the option to select a number of parameters to assist the user to find more relevant information related to their queries, such kind of parameters are considered part of the user profile that is previously stored in the MSDB. And based on the selected parameters the system will rank and return the results to the user.

- **Search Relevance**

To personalize the search results the search relevance parameter will be specified by the user for more accurate and relevant information that match the user query. By selecting this parameter the retrieved document should match the requirement for information and rank the search results based on the user interest. The retrieved documents within this parameter should be relevant to the search query or to the context of the request in which the search was conducted. Moreover, it will also assist the user in a successful conclusion to his search.

Ranking of the results will cover the most relevant results to the search query to the less relevant results. For example, if the user is an optician searching under the subject of (cataracts) the most relevant results for the query could be (caused, symptoms, and treatment), and the less results can be (eye glasses).

- **Search History**

This preference allows ranking the results based on the past search including the keyword entered by the user, the saved URLs, and pages that the user have browsed.

By selecting this parameter the user will be able to select one of the already viewed keywords as well as the visited pages. The search history can be saved in the system when the query is submitted by the user, the system will return a set of results to the user. The user can then select the information to view some results by clicking on them. Now the reviewed information will be saved as a history for past queries and it will be available for the user to view when the user selects the history preferences.

The history parameter is useful for inferring a user's information need, it will assist the user to summarise the previous search using the summary of the past search, rather than re searching using the same query. Moreover, it would be beneficial for the user to select this option when they search for a similar information in the future to save their time.

- **Search Material type**

This preference allows the user to select the type of information source that matches the user's needs. These sources can be books, articles, journals, images, audio, video.

- **Search Recommendation**

This parameter will rank the search results based on other user's recommendation on the items and their feedback it will focus on finding similar items of interest that users recommend it for a similar query. User's feedback on the items will be presented in the form of ratings for items either as value, (relevant, not relevant), or by using a graded system like (not relevant, somewhat relevant or very relevant). In the same field, the user will be able to identify the most relevant item based on this feedback. For example, if the user is a pharmacist searching for (steam inhaler device) and has selected his preference for the search results to be listed based on the recommendation the results will be listed from the highly recommended device to the lowest one that what similar user of interest have recommended it.

5.8.5 Personalization and Ranking of the Search Results

Personalization is a technique used to filter and re-rank search results to present the right information to a particular user at the right time.

The first step in personalization process is gathering the user data through the previously browsed history and the query made by the user, which will be used to provide useful information about the users' preferences. The second step is to rank and filter the results based on the learned profile of the user.

5.8.5.1 Gathering the Information about the User

The user profile is the key element that can be used to collect information about the individual user and to learn about the user interests. User related information will be collected since the user registered his identity with the system for the first time, the user profile will be created and stored on the database. This profile will be updated each time that the user started the session and searched for the information. The first task of a profiling technique is collecting preferences and information for the individual users. Collecting user preferences is a big challenge. Because it is hard for the users to list their preferences and interests in their profiles during the registration. To solve this problem, the system uses alternative ways to collect the user interests to reduce the burden on the user on listing each interest. Therefore, the server will automatically save the search history for the user and the keywords used by the user. In addition to that, the user interest can be collected using a number of terms. An example if the user is interested in searching for (Ibuprofen) this interest can be expressed with terms such as uses, interactions, side effects, clinical data, pregnancy and so on. If the user selected the pregnancy term for multiple time of searches with different related keywords the user profile will show a

high interest in (Ibuprofen and pregnancy) term, as results, the high number of the returned results will be under that term.

In brief, the smart layer will learn about the user explicitly via the information input by the user since the beginning of the session and collecting the user's preferences feedback. It then will learn implicitly by monitoring the user behavioural and capturing the user history by the server via recording and monitoring the user's past search queries and the web pages accessed by the user and the selected results.

When user's information collected the user profile will be created. And because of the purpose of personalization the user profile will be build based on the user-term weight. Which means each term in the user profile will be weighted. The weight of the item represents the important of that team to the user in which a high weight of a term indicates that the term is high significance for the user and a low weight of the term indicates that the term is not important to the user. As an example, suppose there is a user who is interested in "short-sightedness surgery" and has previously used "corneal surgery, retinal surgery, and laser eye surgery". Now, the user profile should have a high weight for the related documents under the subjects of "laser eye surgery" and low weight in the others.

5.8.5.2 Learning Different Users Interests

When the user submits the request and the document returned from the web. The result adaptation engine (RAE) will be invoked by the server to filter the search results and to match the retrieved documents with the user profile. This process aims to achieve the goal of personalization by filtering out the irrelevant information and re-rank the search results based on the user profile.

When the filtering is applied search results should match each user interests. To clarify this information this section will show an example of how the system extracts the user interest and how the smart layer involves to customize the search results.

This example shows two users form medical field who have different area of interest: one is a surgeon and the second is a Paediatrician both are trying to access information about “children and diabetes”. This example try to clarify how different type of users performed the same search and retrieved different results based on the learned profile and how the system returned different information for each of them.

In this example it is assumed that both are searching for the keyword mentioned for the first time using google as a search method, and none of them has recorded any history yet , in this case google may retrieve almost the same results for both users and the results list looks as shown in figure 5.11.

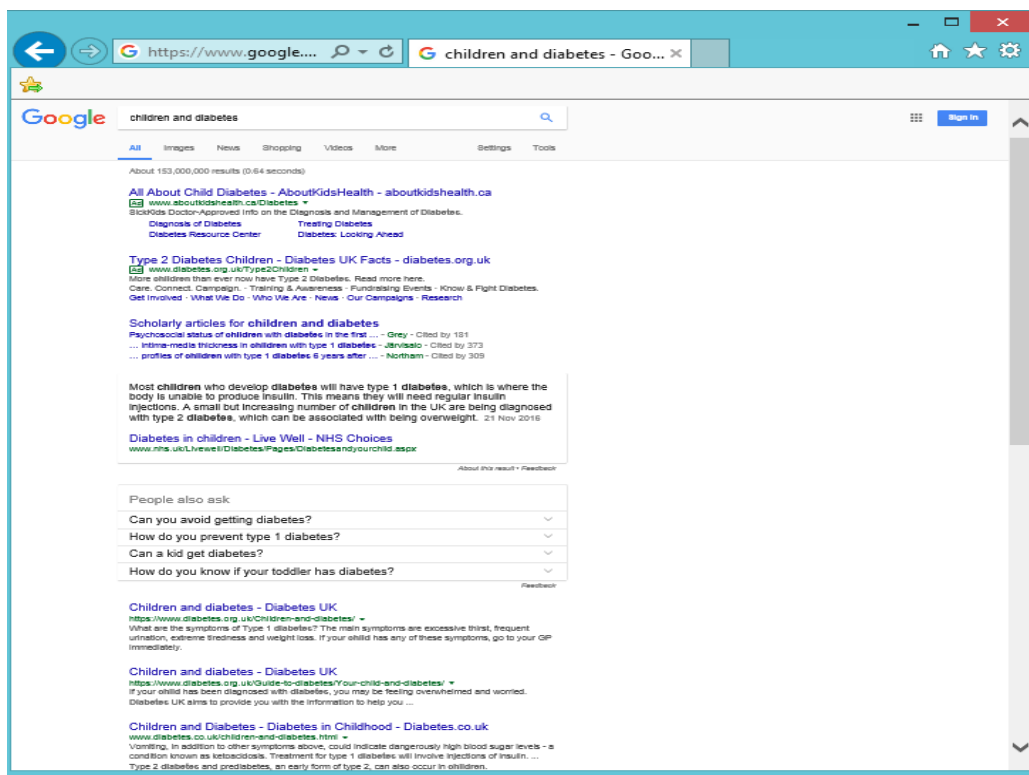


Figure 5. 10 Search Results for Children and Diabetes

Typically, each user will select the relevant information to the query and if not relevant the user may keep browsing and clicking the next page, searching using alternative keywords or changing the information sources. Now once that done each user will select the relevant result and once the user clicks on that information and read the information, this information will be stored as a search browsing history in the user profile, These browsing histories contain the URLs visited by the user and the dates and times of the visits. These browsing histories can be accessed by the user any time later. In case that the user has changed the search keywords or choosed an alternative keyword related to the same subject and match the target of the search, the profile will save the search keywords and build the user preference, at each time when the user changed the keywords. And when both users access the same information in the future the profile will show a high interest of information based on the user interest for example the case of our example the profile of the Paediatrician will show high interest regarding symptoms, medication family awareness while a high interest in diagnostic and blood test and glucose level for the surgeon as showing in figure 5.12 .

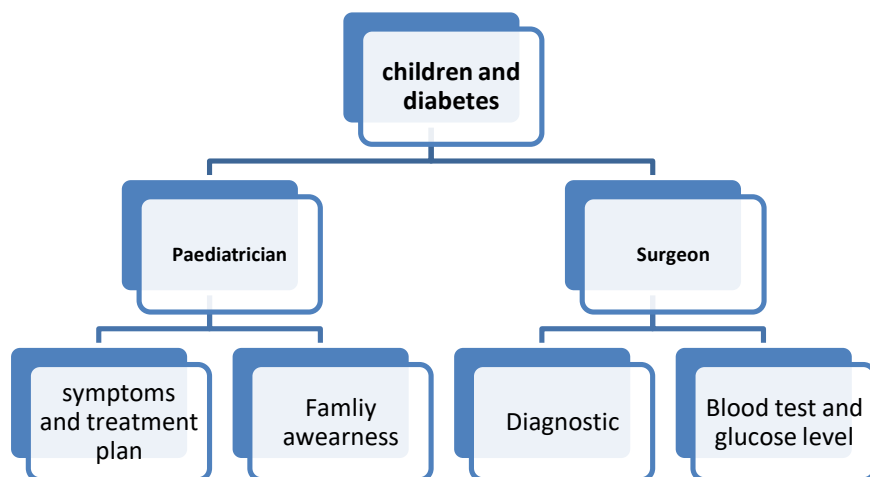


Figure 5.11 User Profile Interest

User profile will present the user interests and it will continues adding any new preferences and update the user information each time that the user access to the system.

When both users have done multiple numbers of sessions searching for the same area, and now both trying to search for the same keyword (children and diabetes) when the server download the user profile it will identify the user preference from the profile and the previous session (the saved history). The system will return different search results for each user some of them may be the same but the majority of the results will match the profile of the user. Table 5.7 shows the expected results that each user may retrieve based on the learned profile.

Table 5.7 Expected Results for the Search Topics

Paediatrician	Surgeon
<ul style="list-style-type: none"> • Symptoms • Investigation and blood test • Insulin • Parents awareness and child's reaction • Physical activity • Health and diet • Monitoring devices • Glycated haemoglobin test • Pregnancy and diabetes • Insulin pump therapy • Fine needle syringe • Insulin pen • Insulin sensitivity • Glucose level test 	<ul style="list-style-type: none"> • Diabetes type1 and 2 • Cause and symptoms • Initial diagnosis • Management and monitoring of glucose level. • Islet cell transplant • HbA1C test • Diabetes's screening test • Prevention • Weigh loos and diabetes • Insulin pump implant • Anaesthetist • NB insulin therapy

5.8.5.3 Ranking the Search Results based on the User Profile

As mentioned earlier that the smart layer aims to improve the relevance of the results by personalizing the results using the user profile

The second step of personalization is ranking the search result in order to achieve the final goal of the smart layer. The collaborative filtering system chosen to filter the search and rank the results based on the user's profile. The method of filtering the results based on the CF is based on the current methods found in the literature such as the method done by Rohini & Ambati, 2(006)., Speretta and Gauch (2005)., Abhinandan et al (2007).

Ranking of the results is done by first retrieving a number of documents from the web matching the user's query. Then filtering the results and ranking them using the collaborative filtering. This section briefly reviews the collaborative filtering algorithms that will be used to rank the results

Overview of the Collaborative Filtering Algorithm

The Collaborative filtering selected to rank the documents based on the similarity between subset of user and the active user. The subset of users is first chosen based on their similarity to the active user, and a weighted combination of their rating is then used to produce predictions for the active user. The CF algorithm is summarized by Kazunari et al.,(2004) as following steps:

- 1- Weight all users with respect to similarity to the active user. This similarity between users is measured as the Pearson correlation coefficient between their rating vectors.
- 2- Select n users that have the highest similarity with the active user. These users form the neighborhood.

- 3- Compute a prediction from a weighted combination of the neighbors rating term weights.

In step 1, $S_{a,u}$, which denotes similarity between users a and u , is computed using the Pearson correlation coefficient defined below:.

$$S_{a,u} = \frac{\sum_{i=1}^I (r_{a,i} - \bar{r}_a) \times (r_{u,i} - \bar{r}_u)}{\sqrt{\sum_{i=1}^I (r_{a,i} - \bar{r}_a)^2 \times \sum_{i=1}^I (r_{u,i} - \bar{r}_u)^2}},$$

Where $r_{a,i}$ is the rating given to item i by user a , and \bar{r}_a is the mean rating given by user a , and I is the total number of items

In step 2, i.e., neighborhood-based methods, a subset of appropriate users is chosen based on their similarity to the active user, and a weighted aggregate of their ratings is used to generate predictions for the active user in the next step 3.

In step 3, predictions are computed as the weighted average of deviations from the neighbors mean:

$$p_{a,i} = \bar{w}_a + \frac{\sum_{u=1}^n (w_{u,i} - \bar{w}_u) \times S_{a,u}}{\sum_{u=1}^n S_{a,u}}$$

where $p_{a,i}$ is the prediction for active user a for weight of term i . $S_{a,u}$ is the similarity between users a and u , and n is the number of users in the neighborhood

5.8.5 Handling Request over the System

Before giving details about how the smart layer handles the data it is important to mention that the framework provides two types of searches based on the user subscription

- First is called the general search which is performed by the ordinary user who has not performed any session, or has a low number of search and has no subscription to any services.
- Second type is called the expert search which is performed by a number of users who is expert in using medical terminologies and have already subscribed to many services and performs a number of sessions.

As motioned earlier that the smart layer created to customise the search results in order to avoid the huge amount of the retrieved results. However, this layer will not provide any support to the general user or the user who just registered to the system and have done just one session. Therefore, this layer will be limited on the expert users only.

The performance of this unit is based on the decision made by the S-MIC. The S-MICF will contact the MSDB to download the user profile. Once the server has downloaded the user profile and has identified the user type and subscription. It will then decide whether to invoke the smart layer or not. If the user profile shows that the user has registered and set the first session, in this case, the smart layer will not be invoked, the server then will perform a general search process. If the user profile shows a subscription for the expert user, now the smart layer will be invoked in order to customise the results.

The data plan description for the expert search and the call flow for the general search are presented in chapter 4.

5.9 Summary

A description of the smart layer and the steps followed in order to design this layer is presented in this chapter. The chapter presents a case study to investigate the number of results obtained from 4 different databases. An experiment was carried out by 4 groups of users from different medical fields searching under their keyword. The experiment presents each group search results and then identified the gap and the need for a system that has the ability to customise the search result based on the user profile. The aim and the design of the smart layer are defined. A description of the smart layer components along with the search parameters is presented. The steps of ranking the search results are covered. The chapter concludes with a description of handling the request over the smart layer. The next chapter will evaluate the framework and will test the validity of the smart layer.

Chapter 6: Framework Evaluation

In Chapter 3 the overall framework design was agreed after thorough investigation of the system requirements and needs. It then presents the layered model along with its functional entries. While chapter 4 gives a detailed description of the signalling plan, the flow process of the messages and describes the two scenario for the system subscribers. In this chapter, the designed framework will be validated via a simulation on a discrete event simulator for estimating the performance measures such as the delay and throughput under heavy loads. The evaluation is based on number of assumptions will be mentioned later on in this chapter, the evaluation will take into a consideration the response time of the request and packet loss ratio which reflect the scalability and reliability for the system.

6.1 Evaluation Methodology

6.1.1 Databases Performance Evaluation

Performance evaluation of databases is crucial to better understand the performance metrics related to it. There are different models to evaluate and validate database performance, it can be summarized as follows:

❖ Queuing Models

The operational nature of databases can be simply characterized as an interconnected queues, in which transactions are represented as concurrent jobs served for a specific amount of time. Each queue represents a separate database with the possibility of having multiple queues that are interconnected and distributed over the network.

❖ **Cost Models**

The physical related performance aspects are best represented with cost models in such models the database response time and the performance is described per each database and compared against other accordingly.

❖ **Simulation Models**

Simulation is the process of approximating the behavior of a system over a period of time. The simulation model is used to gather data as an estimate of the true performance characteristics of the system.

Simulations gives near real estimate of physically deployed systems. The interconnections with other interfaces along with the hardware specifications are taken into account which may be missed by other models. It also has the ability to merge both the hardware and software packages in one setup by defining the process and application instances on top of the components and hardware entities.

❖ **Benchmarking Analysis**

Having a baseline performance and comparing it against other system performances using benchmarking method is one of database performance evaluation techniques. The differences in implementation techniques in addition to the introduced workload (stress) over databases create different system responses that reflect performance variations. Such variations can be exploited to best benchmark the systems and decide the performance metrics for each one of them.

In this study, a simulation model of new framework for digital medical libraries along with a benchmarking technique are introduced to best evaluate the performance of medical database libraries. Benchmarking technique requires that the systems be implemented, so first the system implemented by running a simulation model over the OPENIT simulator and then by running the workload on the system the configurations of the system will supply information which can be used to compare and evaluate the results. These two techniques chosen because both provide the most valid performance results.

An abstract model for the simulation is adopted as shown in figure 6.1. Without loss of integrity, the basic functions of the proposed framework which already described in chapter 3 for the following entities (Access layer UE, Control layer P-MICF, S-MICF, HSDB, SL, and the Application layer) are represented in the abstract model. The smart layer (SL) entities are merged as part of one single cache server for evaluation purposes, in which the server will be handling both the requests and responses into and out of the smart layer.

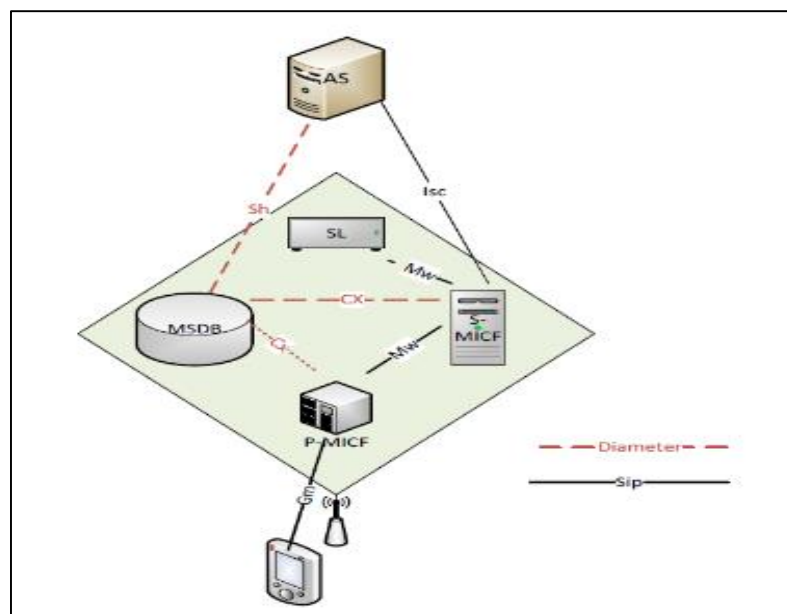


Figure 6.1 Abstract Model

6.1.2 OPNET

The proposed framework implemented on OPNET modeller for academic and research purpose that is licensed for the centre for communication systems research, University of South Wales.

OPNET Network simulator is a tool to simulate the behaviour and performance of any type of network. The version used for the experiment was 9.0.A. The engine of the OPNET Modeler is a finite state machine model in combination with an analytical model. It can simulate protocols, devices and behaviours with about 400 special-purpose modelling functions (OPNET, 2018).

The Optimized Network Engineering Tools (OPNET) is a commercial simulation product of the MIL3 Company of Arlington, VA. It employs a Discrete Event Simulation approach that allows large numbers of closely-spaced events in a sizable network to be represented accurately and efficiently. OPNET uses a modeling approach where networks are built of components interconnected by perfect links that can be degraded at will. Each component's behaviour is modeled as a state-transition diagram. The process that takes place in each state is described by a program in the C language. This makes the OPNET-based models relatively easy to port to other modeling environments. This family of models is compatible with OPNET from version

A detailed description of the OPNET environment including the routers, Subnets, LAN Subnets, node layer, host, packet streams, session sender and receiver are defined in RFC 2490 by Pullen et al. (1999).

6.1.3 Jitter Delay

Definition:

The absolute value of the difference between the Forwarding Delay of two consecutive received packets belonging to the same stream (Poretsky, S et al., 2009)

Discussion

Jitter is defined as a variation in the delay of received packets. At the sending side, packets are sent in a continuous stream with the packets spaced evenly apart. Due to network congestion, improper queuing, or configuration errors, this steady stream can become lumpy, or the delay between each packet can vary instead of remaining constant (Poretsky, S et al., 2009).

Figure 6.2 illustrates how a steady stream of packets is handled .

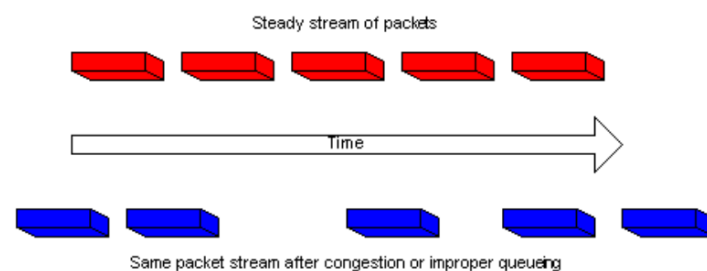


Figure 6.2 Steady Stream of Packets

When a router receives a Real-Time Protocol (RTP) audio stream for Voice over IP (VoIP), it must compensate for the jitter that is encountered. The mechanism that handles this function is the playout delay buffer. The playout delay buffer must buffer these packets and then play them out in a steady stream to the digital signal processors (DSPs) to be converted back to an analog audio stream. The playout delay buffer is also sometimes referred to as the de-jitter buffer. Figure 6.3 illustrates how jitter is handled .

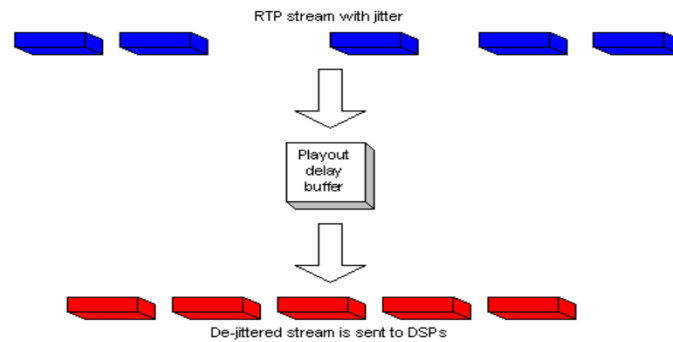


Figure 6.3 Handling Jitter

If the jitter is so large that it causes packets to be received out of the range of this buffer, the out-of-range packets are discarded and dropouts are heard in the audio. For losses as small as one packet, the DSP interpolates what it thinks the audio should be and no problem is audible. When jitter exceeds what the DSP can do to make up for the missing packets, audio problems are heard. Figure 6.4 illustrates how excessive jitter is handled (Poretsky, S et al., 2009).

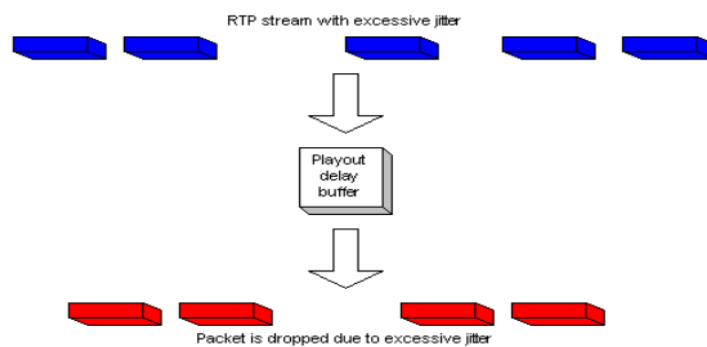


Figure 6.4 Handling Excessive Jitter

6.2 Case Study (User Classification)

When it comes to medical databases, the performance evaluation technique highly depends on the workload generated by the user. In order to understand the nature of the medical databases users and the way they used the database to search for the information. A survey was conducted at Heath Hospital in Cardiff, the medical and emergency staff members were very interested in having a system with certain requirements that are considered essential for their daily job routine tasks. It was confirmed by 43 out of 73 physicians that having a reliable system that allows its user to access medical databases through their devices (Mobile-PCs) with a reasonable response time is of great importance in the medical field as it is usually considered a critical system by nature and usually the delays affect human lives.

The case study of different types of medical staff users at (Heath Hospital) along with different service requirements are presented in chapter 5.

The case study highlighted the following information that helped in building the basics assumption for the simulation model:

- According to the case study presented in chapter 5, the responsiveness of the intended system was evaluated via a questionnaire to determine who will be using the query service among the staff, when the query responses will be needed, and for how long the staff will be able to wait for the results to be returned back, Finally, which targeted database they prefer to use. According to the answered questions (Who, When, How long, and which), the case study classified the intended service into two broad categories;
 - Real-time Detailed Service for Expert Users
 - Non-Real-time General Service for Common users

Both will be referred to RDSE and NRGSC respectively. Table 6.1 shows the different participating groups and the classification that best fit each one of them according to the aforementioned criteria.

Table 6.1 User Classifications

Group name	Sample Size	Occupation	RDSE	NRGSC
Group 1	13	Nurses	√	
Group 2	10	Dentists		√
Group 3	9	Pharmacists	√	
Group 4	11	Medical student	√	

According to the results presented in the previous chapter in figures (5.6, 5.7, 5.8, and 5.9) the keyword search results for the picked sample varies depending on the database used, the staff member expertise, and the keywords applied. This simply shows the differences among the user groups and highlights the need for an abstract layer that customised the search queries according to the user profile that is saved in the system.

6.3 Simulation Model

Simulation model of real time systems simplifies systems evaluation, but when it comes to real time databases access and real time signaling protocols, having an accurate model with realistic assumption and configurations is of great challenge. System workload for example, need to reflect real world traffic to get better estimate of the system performance under different operation conditions.

Real time databases are critical in nature, the need for having a medical system is crucial to send and retrieve the queries within strict time limits to ensure data freshness and system responses especially if the information needed during the patients consultancy or even at the emergency unit. Average response time and throughput are among the most important metrics in database systems. The previously proposed framework is implemented in a hospital site in OPNET as shown in figure 6.5. There are many assumptions made to replicate the environment of the hospital presented in the case study.

The following assumptions were made and configured in the simulation model:

- 1- There are two pairs of medical staff members who are using voice service within the hospital. One of the callers has higher priority than the other one. Caller 1 initiate an interactive voice service to Callee 1 using SIP signalling. Caller 2 initiates excellent effort voice service to Callee 2 over SIP signalling as well. The calls generated are concurrent with exponential generation and uniform repetition distributions. The proxy server is located within the hospital domain.
- 2- There is a subnet with 200 users over Ethernet, each user access a webpage that is connected to a page content server and a video archive server. This assumption is based on the type of services that medical staff members are using in the hospitals nowadays. There are many configurations made to reflect the realistic nature of the webpage content. The video archive and page retrieval servers are located outside the hospital domain.
- 3- The smart layer functionality presented in the framework previously is implemented as a part of a cache server. The cache server gets the whole traffic and decides if it

needs to be forwarded to the database server, page retrieval server, or the video archive server. It will send the response to the server to forward it to the user.

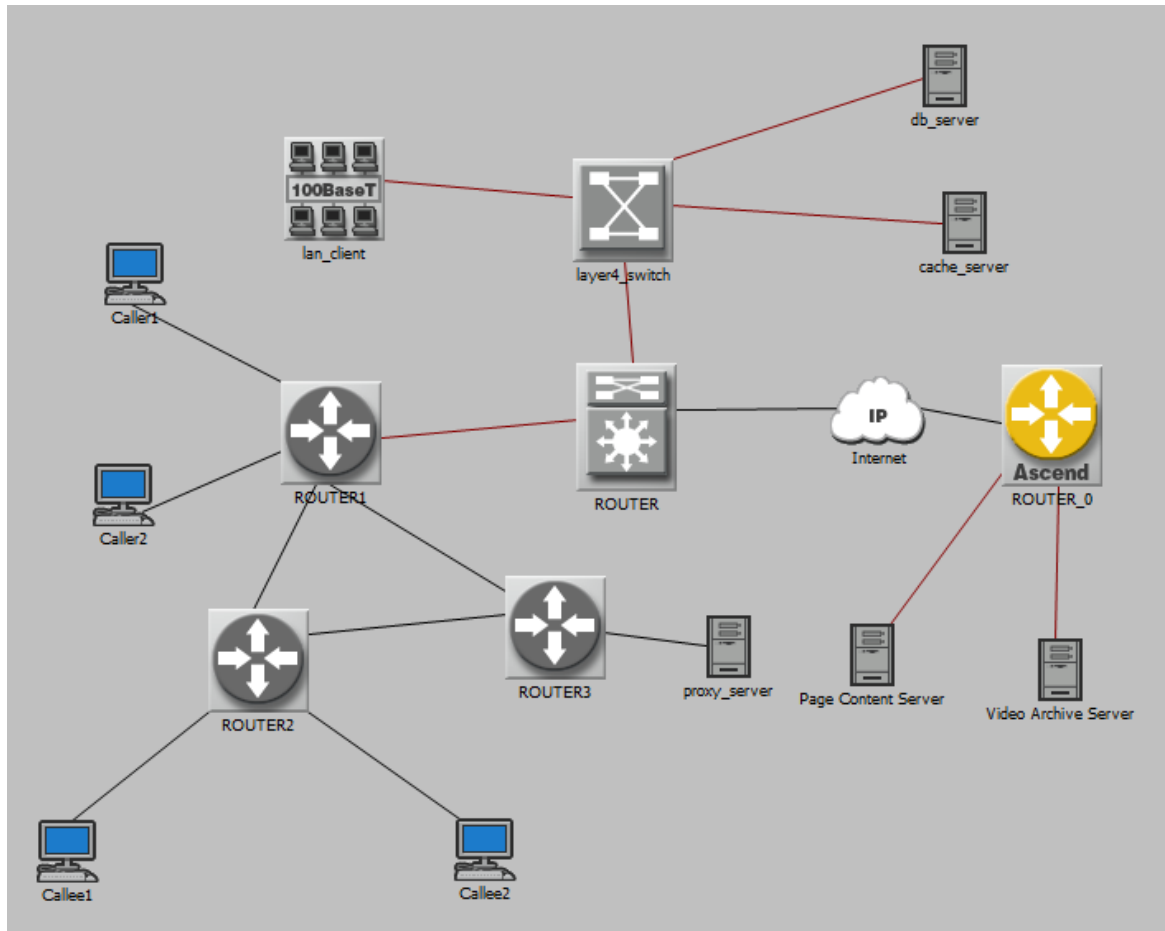


Figure 6.5 OPNET Simulation Model

Figure 6.5 represents the simulation model for the implemented framework which shows 2 users (Caller 1, Caller2) connected to a router within the hospital domain, the router connected to a set of routers for connectivity purposes. Caller 1 is communicating with Callee 1 and caller 2 is communicating with Callee 2.

As mentioned earlier in the assumption that caller 1 is higher priority than caller2 and both are communicating with the other users at the same time.

In this communication there is 2 types of data:

- **SIP signalling**

The communication between the users will not go directly from caller 1, 2 to Callee 1, 2. It will be a SIP signalling INVITE message that sent by the user and goes through the proxy server first for authentication purposes and when the proxy server has identified the user it will send a 200 Ok message to allow caller 1, 2 to communicate with Callee 1, 2 the user first will replay with Ok to the proxy server and then can continue the conversation.

- **Voice data communication**

After the authentication procedure has successfully done by the proxy server, the data voice will go directly from Caller 1,2 to Callee 1,2 without passing the proxy server for the second time, this voice service will be within the hospital domain for example a user at the Xry unit needs to contact a user at the emergency unit

The simulation model also shows a (lan_client) which is a subnet consists of 200 users, database, and cash server (which performs the smart layer functions). The page content server and video archive sever both are part of the web browsing HTTP and located outside the hospital domain. The Swither and the router are connectivity units between the devices.

It is important to mention that in this model we are interested mainly in the response time of the requests and the packet loss ratio which reflects the scalability and reliability concerns for the overall system.

6.4 Results and Discussion

The initial simulation results showed that increasing the load of requests from the hospital domain by sending SIP requests to the MICF server which is supposed to process the requests based on the proposed framework. The response time was increasing with the generated SIP and Voice traffic load. This has a negative impact on both the Quality of Service (QoS) and Quality of Experience (QoE) for the users.

Medical Staff members have shown during the survey that the response time is of great importance especially if the system supposed to serve patients during their consultancy and also in the emergency room where usually the requested task is delay intolerant. And for such system, both the delay and packet loss ration reflect two major requirements such as the scalability and reliability of the system.

Figure 6.6 shows the Cumulative Distribution Function (CDF) that presents the probability of the two caller pairs (with different priority values) the caller with higher priority has a lower end to end voice traffic delay compared with the lower priority calling pair.

The framework deployed Weighted Fair Queuing (WFQ) mechanism based on the user definition. Users with higher priority are the ones who are expected to be less tolerant to network delays, they tend to use real time interactive services.

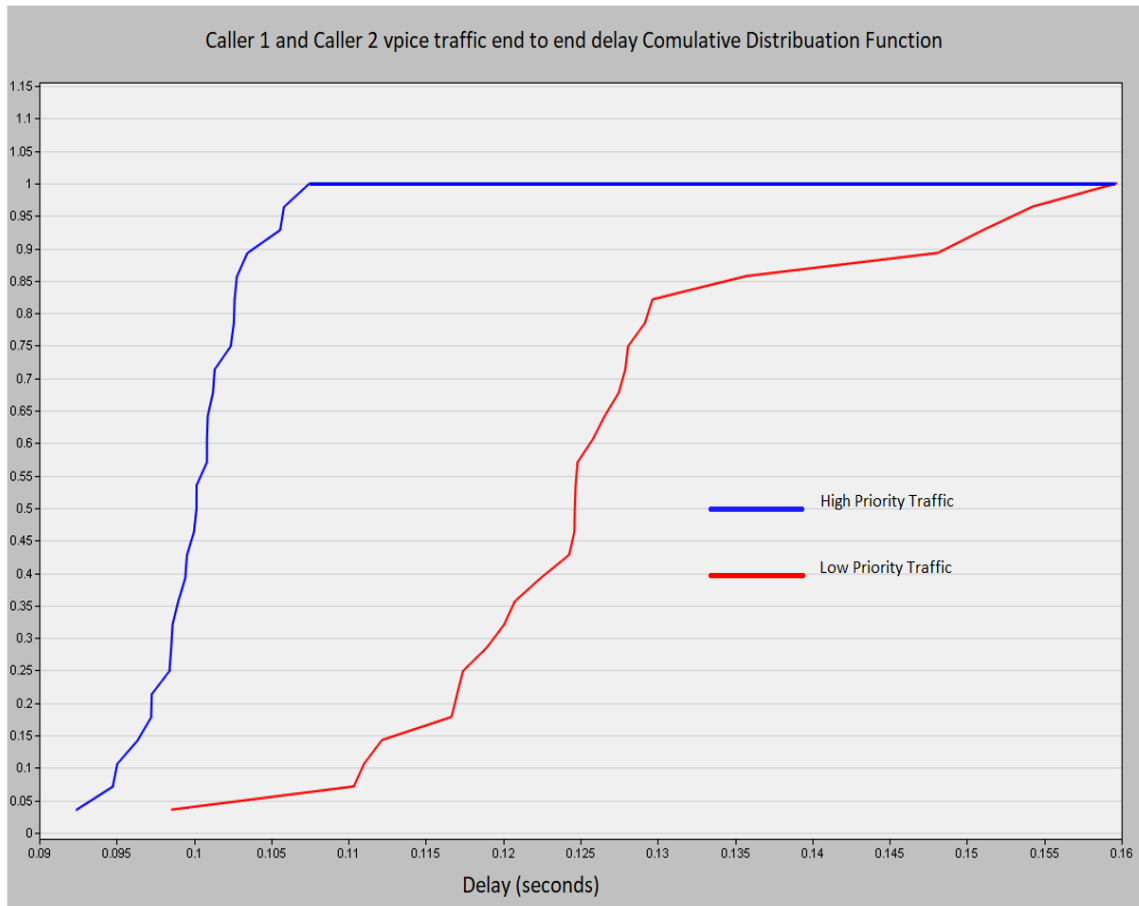


Figure 6.6 End to End Delay CDF

Figure 6.7 shows the voice traffic delay variations (Jitter delay) of the transmitted voice packets. Although the voice coding scheme and the network topology play important rule in the Jitter delay, both factors were the same for the two callers and as a result we get less jitter delay values for the higher priority traffic compared with the lower priority traffic. The jitter delay contributes negatively to the audio quality.

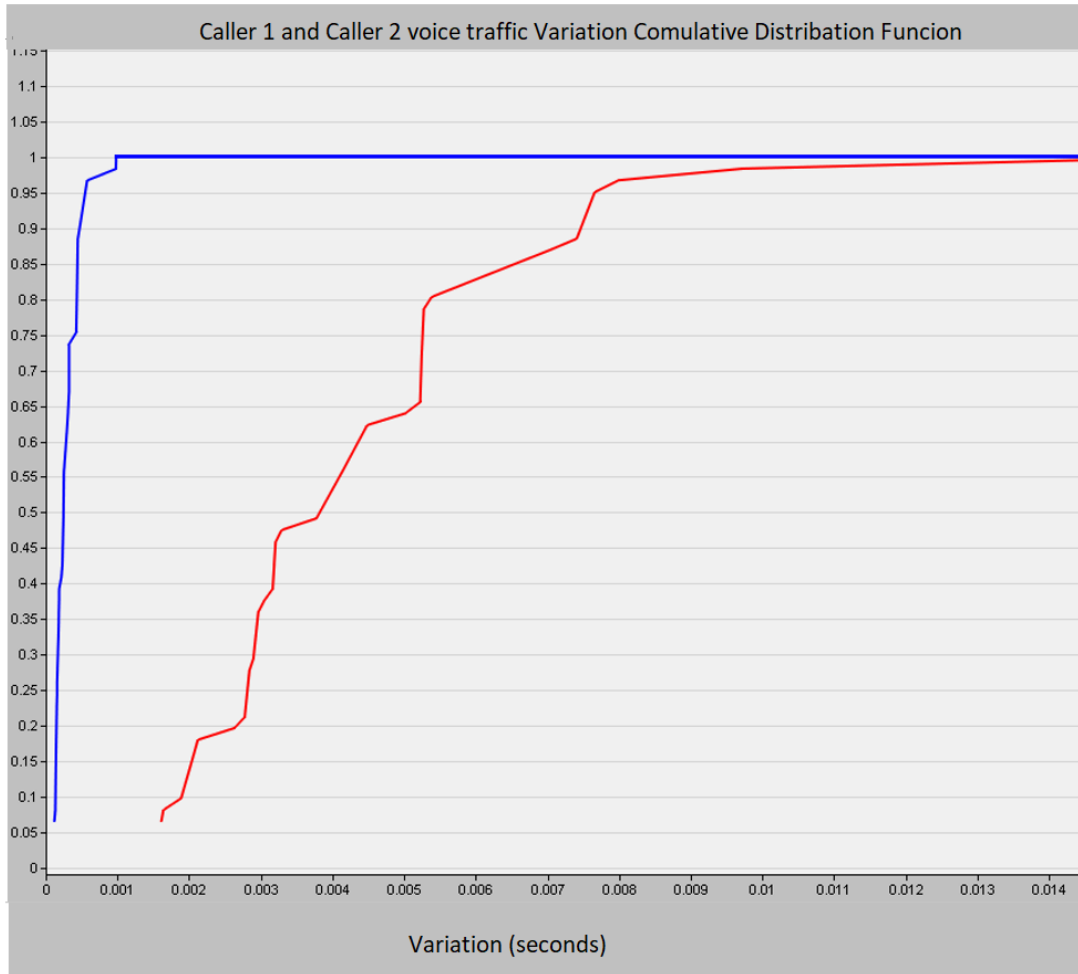


Figure 6.7 Voice Traffic Delay Variation CDF

This experiment was conducted to investigate the way that medical staff members use interactive services and search their information and the number of results they obtained from different preferred methods. The outlined step is done in order to identify a way to customize and reduce the amount of results returned and to provide a filtering method to avoid a large number of irrelevant information. This experiment will outlines why the smart layer designed as is described previously in chapter 5. The smart layer was deployed using a cache server in which the hit rate was varied and the requests were forwarded to the destination servers according to the hit rate value.

Figure 6.8 shows the CDF of the HTTP requests throughput. The throughput is an indication of remote page retrieval system load. The higher the throughput the more delays will be experienced as more requests need to travel further. The smart functionality of the cache server reduces the need to send remote requests by recognizing the user profile and retrieving the information from cached databases or from a dedicated nearby servers, this creates the distinction between the general search and the more customized expert search.

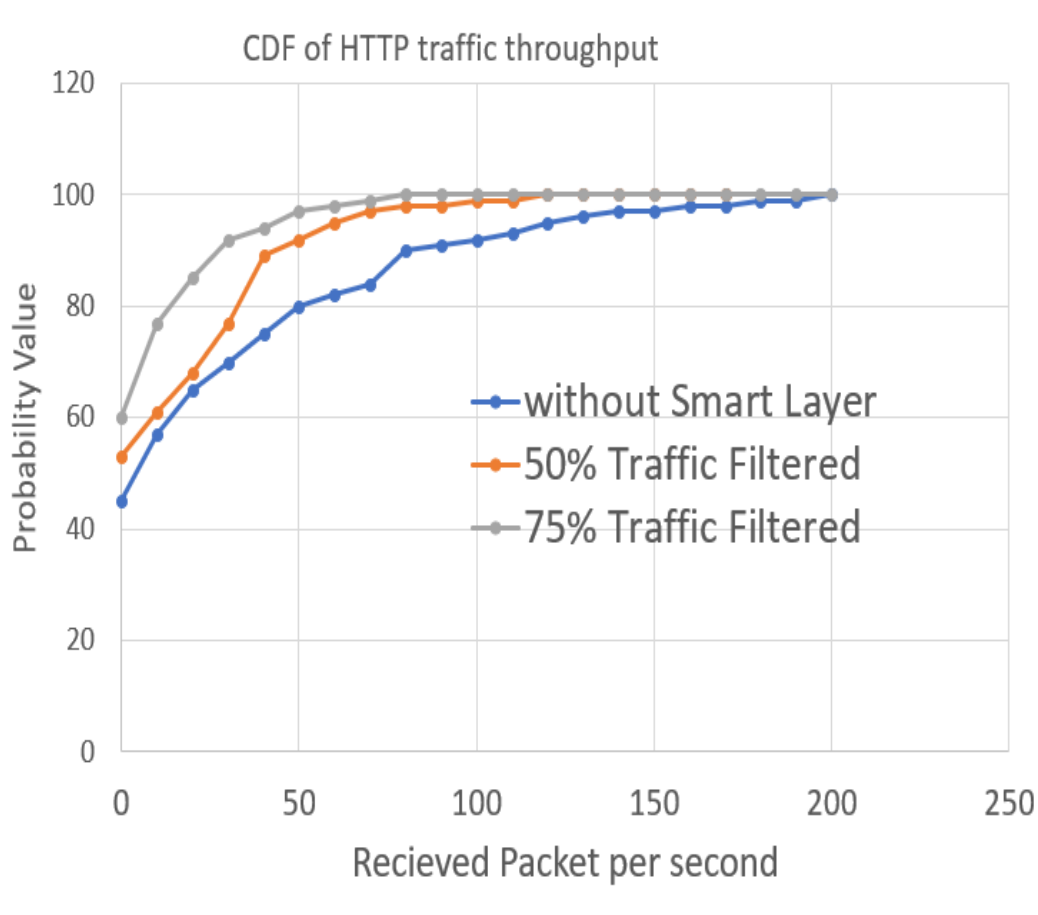


Figure 6.8 HTTP Request Throughput CDF

The traffic throughput showing in figure 6.8 presents three scenarios:

- **Scenario 1:** Not using the smart layer (all traffic is forwarded to the server)

In this scenario the user request will be sent to the final destination which the application programming interface (API) and the results will back to the user in the same way via the MICF server with no involvement from the smart layer. The grey line in figure 6.8 presents the traffic throughput for scenario1 (not using the smart layer) which is very low comparing to the others and that caused a huge delay for the request.

The data flow for the request for scenario 1 is presented in figure 6.9.

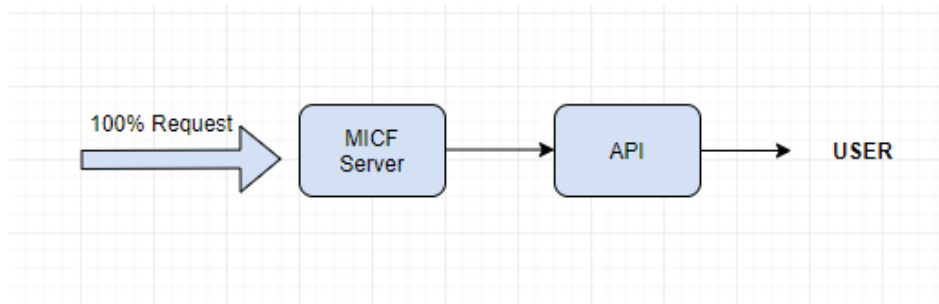


Figure 6.9 Scenario 1 No Traffic Filtered

- **Scenario 2:** 50% Traffic filtered

Sending 50 percent of the traffic and processing locally the remaining part

In this scenario the user will be identified by the server then will send a 50% of the traffic to the cash server which performs the Smart layer filtering function and will

process locally the remaining part of the traffic user request. The data flow for the request is shown in figure 6.10.

The red line in figure 6.8 presents the traffic throughput for scenario2 (processing 50 traffic filtered) which is still low comparing to scenario 3. And that also caused a delay for the request.

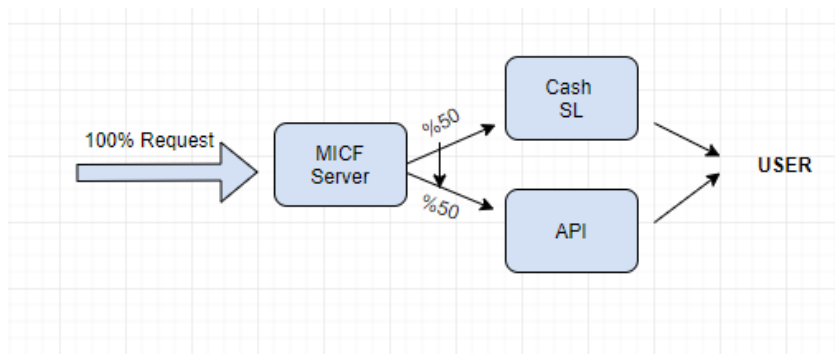


Figure 6.10 Scenario 2 50% Traffic Filtered

- **Scenario 3: 75% Traffic filtered**

Sending 75 percent of the received traffic and processing locally the remaining part. In this scenario the user will be identified by the server then will send a 75% of the traffic to the cash server which performs the Smart layer filtering function and will process locally the remaining part of the traffic user request. The data flow for the request is shown in figure 6.11.

The blue line in figure 6.8 presents the traffic throughput for scenario3 (processing 75 traffic filtered) and that clearly showed a high throughput comparing to the other scenarios and that shows less delay.

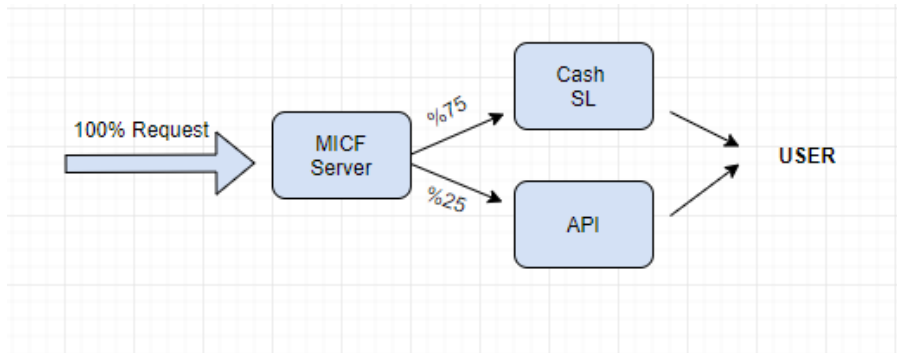


Figure 6.11 Scenario 3 75% Traffic Filtered

From the three scenarios, it can be summarised that eliminating the smart layer functionality will increase the overhead over the remote servers and vice versa. Similarly as shown in figure 6.12, the processing time of remote requests is enhanced by filtering part of the aggregated traffic at the client side.

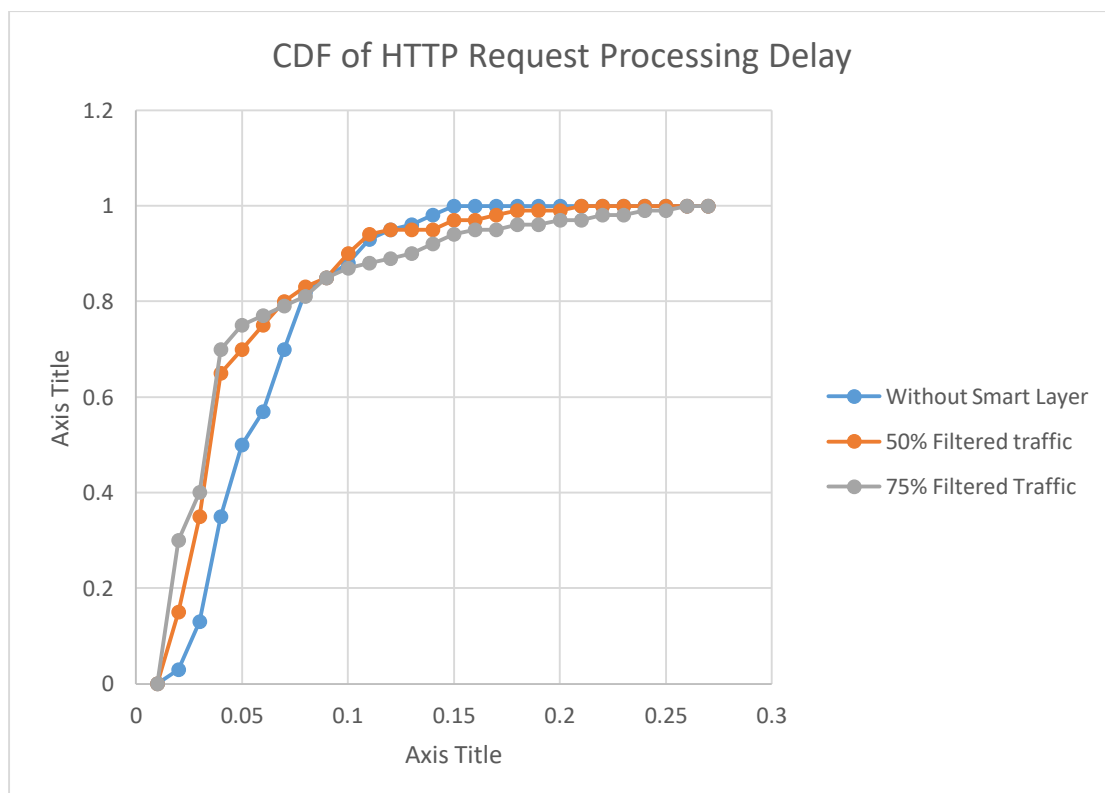


Figure 6.12 HTTP Requests Processing Delay

6.5 Summery

Medical services are critical in operation by nature, the response time of the access systems along with end to end delay of the real-time services imposes challenges to the quality of the running services. Using the IMS functionality along with other IP priority classification mechanisms improve the system responsiveness and scalability and robustness. Moreover, using SL functionality improved the system performance and showed less delay of the response time and that proves the role that the smart layer plays and the value that adds to the framework. It was shown that the proposed framework has improved the performance of the overall system considerably.

In this chapter the designed framework has validated via simulation model using OPNET simulator, the connectivity and traffic workload is tested by users from medical domain, the smart layer functionality is tested by running three scenarios of traffic. The simulation model confirmed that the SL functionality has improved the performance of the framework.

Chapter 7: Conclusions and Future Research

Researchers in medical information retrieval have proposed a number of systems that contribute significantly to solving the problem of finding relevant medical information on the web and helping the physicians to make a decision and treatment plan for their patients. Chapter 2 reviewed a variety of methods for finding out and accessing the medical information. These methods are well designed but did not focus on the retrieved results and how to personalise the results in order to get a fewer number of results to satisfy the user's needs. Therefore this research presents a solution focused on customising the results based on the user profile. This chapter will draw together the recommendations of this thesis and will suggest directions for future research.

7.1 Conclusions

This thesis has discussed the medical information retrieval system and presented a number of methods that been used by users from the medical field to retrieve medical information. The thesis also identified a number of problems that user faced while using these methods to locate their information, which lead to the key contraption to this work. A framework for medical information retrieval system (MICF) for retrieving relevant information based on user profile has been introduced in chapter 3. The framework contains four main layers and each layer contains a number of components to perform a number of tasks in the proposed system. While Chapter 5 introduced the smart layer which plays the main role of customising

the search results to each user based on their profile. This step made in order to avoid the significant amount of the returned results from the system.

The ability of the system on handling the user request has been defined through a number of a processes which include the registration process, the authentication procedure and the process of the INVITE message. The call flow for the mentioned processes are presented in chapter 4. The need of the smart layer is proved by a number of steps detailed in chapter 5 and then a proposed framework has been evaluated via a simulation model presented in chapter 6.

7.2 Achievements

The aim of this thesis is to present a comprehensive framework for medical information which has ability to enhance the search results customization based on the user profile.

The first issue of this thesis is designing the framework. The thesis has focused first on studying the existing medical information retrieval system and identified a variety of methods used by physicians to access the medical information and find the advantages and disadvantages of them, and then propose and design the layer model architecture . The framework adapted the horizontal layer architecture from the IMS and designed as follows:

- First step of the framework is building the layers and identifying the functions. This step has focused on the technical part of the framework which includes defining the framework entities along with their functions (Chapter3).
- Second step is defining a number of protocols and interfaces which are selected from the existing system and adapt them to fit in our system for the purpose of connectivity between each unit in system (Chapter 3).

- As the previous step has identified the protocols, this step focused on how these protocols and interface being used to handle the user request. This stage contains the following steps :
 - Description of the SIP message and the session description protocols.
 - Detailed description of the REGISTER message clarified with call flow for the process.
 - Description of how the system performs the authentication along with the call flow of the process.
 - Description of handling the user INVITE request.
 - Two different scenarios for two types of users the general user and the expert user are described along with the data flow for each user.

The second issues of this thesis is to address the search results customisation in order to solve the problem of retrieving large amount of results. For that purpose the thesis has introduced the smart layer as a solution to adapt and customise the search results in order to improve the relevance of results at lower number based on each user' profile. To test the validity of the smart layer the thesis has presented an experiment, performed by a number of users from the medical field and contains a number of steps covered in chapter 5. From these experiments the most outstanding findings are:

- The experiment carried out by different groups of users from medical field each group contains of a number of users within the same medical specialism. The experiment shows that users in each group has a different area of interest. Therefore the search results obtained from different methods were different in content and number.

- The search methods used by the participants to perform the experiment returned a significant amount of results for all groups, and that can be an evident of the needs of improvement on the medical information retrieval in order to consider the search result customization to avoid the aforementioned problem.
- The expert user have confirmed the role that the smart layer would play on the search personalization. The simulation model presented in chapter 6 has proven the validity of the framework.

7.3 Limitations

Although the educational background of the author is in the library and information science, the area of computing science and the subject under health care was a new experience, therefore this thesis required a huge effort for the author especially for designing the framework, understanding IMS communication, setting the protocols and defining the data plan for the request and understanding and using the OPNET simulator. However, searching for the subject of medical information retrieval system has been interesting and valuable not only for system design purposes but also for extending the personal knowledge.

The author also has gained an overview of the health-related information retrieval systems used by the clinics including search engines, databases and mobile applications.

While this study has presented a framework to address the problem of search results customization and presented the smart layer as a solution to customize and minimize the search results based on the user profile. There is a limitations on the solution, the limitations are summarised below and they provide a lead for the future work to be done.

- **Implementation**

The designed framework was limited only to the control and protocol level, which includes the connectivity and the protocols, without focusing on the application level and designing an interfaces for services.

- **Evaluation**

The evaluation focused on the request overload to the MICF server in order to identify the delay of the response time of the requests and also in order to identify the smart layer function and did not includes other type of performance evaluation for the databases .

- **Smart layer design**

- This study aimed to involve a variety number of physicians from different specialisms, but because of the time limitation of the physicians this study was limited to four groups only from the medical field.

- The experiment performed by users was limited to four sources of information although that some of them contain many generic biomedical information such as statistical, figures, images and other unrelated medical information.

- Patients would show the fact of how the system is designed to adapt different types of users and how to support simple patients' request, but due to the lack of satisfactory interaction from the patients to participate in our experiment, the study did not involve patients.

- The smart layer designed for the purpose of ranking, customising the search results based on user' profile. Therefore, the thesis focused on the designing of the smart layer and setting a number of features for the purpose of customising the search results. These features were limited to a number of

parameters specified in chapter 5, these parameters selected to perform the tasks of which the smart layer designed for. Also because the focus of the smart layer was to provide the user with relevant information based the learned profile, other additional feature that can improve the performance was not included.

7.4 Future Work

This study has raised several issues apart from its focus area and limitation. The study listed here as a potential for future research. With respect to the research the study suggests three directions of future research.

- **The implementation**

The implementation considered the medical users, and the needs of having such system that can deliver the results as fast as possible to the user, for that purpose the system focused on the response time to the user request to identify the delay and that helped to investigate the role that the smart layer plays on customising the search result. However, it would be beneficial for the researcher to design an API for the services.

- **User feedback**

It would be beneficial to consider the user feedback to improve the retrieval performance. User feedback is an important tool to improve the system performance and to investigate the strength and weakness of the system. of course it will require an extra effort from the user but it would be interesting to investigate the Physicians feedback on the system on two stages first during patient consultants and second in their daily life.

Another interesting point would be to evaluate the system performance including the relevance, accuracy and efficiency.

Another point on the evaluation is to extend the evaluation of the system and to consider the patient evaluation. As this study has evaluated the system MICF based on the physicians' assessments. Another interesting approach would be to perform the patients' evaluation in which patients actually use MICF to seek their medical information on their daily life, a more reliable assessment of the system performance might be obtained.

In addition to the suggestions for future research stated above, the study suggests three design issues that may be improved by future development.

1. The system might be improved by considering more investigation of the physicians' needs. In order to be efficient on designing the application layer it would be helpful to address the user needs and the most services they would use to access the medical information. This step can be done using initial survey to determine the services. The literature in chapter 2 has covered the user' needs mentioned in a number of studies. It would be interesting to the author to obtain more information regarding the physicians' needs from physicians themselves by conducting a study to determine their needs.

It would be beneficial to improve the system MICF to handle multiple languages and terminology mapping. However, since terminology mapping and language mapping were not the aim of the current research, the study suggests considering specific medical ontologies either MeSH terms or UMLS terms for mapping the medical terminology.

Medical Images retrieval system is a different area if information retrieval in the medical domain. Did not cover in this thesis but in order to provide multimedia services including the medical images and x-ray it would be interesting to do further research on the medical image retrieval.

Lastly, Because of that, the stand-alone system and the patient records system were out of the scope of this thesis it would be interesting to do further research on them for personal knowledge.

References

- Abhinandan S. Das, Mayur Datar, Ashutosh Garg, and Shyam Rajaram. (2007). Google news personalization: scalable online collaborative filtering. In *Proceedings of the 16th international conference on World Wide Web (WWW '07)*. ACM, New York, NY, USA, 271-280.
- Abroms LC., Padmanabhan N., Thaweethai L., and Phillips T (2011) iPhone apps for smoking cessation: a content analysis. *Am J Prev Med* **40**(3), 279- 285.
- Ajuwon GA., & Popoola SO. (2014). Influence of internet accessibility and demographic factors on utilization web-based health information resources by resident doctors in Nigeria. *Afr J Med Sci* **43**(1), 61-71.
- Alghoson, A. M. (2014) Medical Document Classification Based on MeSH. 2014 47th Hawaii International Conference on System Sciences, 6-9 Jan. 2014. 2571-2575.
- Altschuler J., Margolius D., Bodenheimer T., & Grumbach K. (2012) Estimating a reasonable patient panel size for primary care physicians with team-based task delegation. *Annals of Family Medicine* **10**(5), 396-400.
- Alvarez-Cavazos, F., Garcia-Sanchez, R., Garza-Salazara D., Lavariega, J. C., Gomez, L. G. & Sordia, M. (2005). Universal access architecture for digital libraries. *Proceedings of the 2005 conference of the Centre for Advanced Studies on Collaborative research*. Toronto, Ontario, Canada: IBM Press, pp. 12-28
- Amsbary JH., & Powell L. (2013) Content analysis of self-help internet discussion boards: What they do, what they don't do. *Human Communication: A Publication of the Pacific and Asian Communication Association* **16**(2), 49-57.
- Andrews JE., Pearce KA., Ireson C., & Love MM. (2005) Information seeking behaviours of practitioners in a primary care practice-based research network (PBRN). *Journal of Medical Library Association* **93**(2), 206-212.
- A. Niemi., Arkko. Torvinen. (2002), Hypertext Transfer Protocol (HTTP) Digest Authentication Using Authentication and Key Agreement (AKA).RFC 3310. Network Working Group Request for Comments.
- B. Aboba, P. Calhoun, S. Glass, C. Perkins, B. Patil, D. Mitton, G. Zorn. (2000).Criteria for Evaluating AAA Protocols for Network Access, RFC 2989, November2000, URL: <http://www.ietf.org/rfc/rfc2989.txt>. (2001-11-08)

- Barghout F., Halasey L., Said T., Mousa AH., & Dabdoub A. (2009) Evidence-based medicine among Jordanian family physicians: Awareness, attitude and knowledge. *Canadian Family Physician* **55**, e6-13.
- Bates BR., Romina S., Ahmed R., & Hopson, D. (2006) The effect of source credibility on consumers' perceptions of the quality of health information on the Internet. *Medical Informatics & the Internet in Medicine*, **31**(1), 45-52.
- Boonstra A., & Broekhuis M. (2010) Barriers to the acceptance of electronic medical records by physicians from systematic review to taxonomy and interventions. *BMC Health Service Research* **10** (231). Doi: 10.1186/1472-6963-10-231
- Brandt A. (2012) A reader's guide to 200 years of the New England Journal of Medicine. *New England Journal of Medicine* **366**(1), 1-7.
- Brusilovsky, P., Kosba, A. and Nejdl, W. [2007], The adaptive web: methods and strategies of web personalization, Lecture notes in computer science, Springer, Berlin
- Budiu R., & Nielsen J. (2011) *Usability of iPad apps and websites: Research findings a year after launch*. Fremont CA: Nielsen Norman Group.
- Bukachi F., & Pakenham-Walsh N. (2007) Information technology for health in developing countries. *Chest* **132**, 1624-1630.
- Bryant SL. (2004) Information needs and information seeking behaviour of doctors. *Health Information and Library Journal* **21**, 84-93.
- Carney, S. & Koufogiannakis, D. & Ryan, P(2004) "Library Services for Users of Personal Digital Assistants: A Needs Assessment and Program Evaluation." *portal: Libraries and the Academy*, **4** , 393-406.
- Can, A. B. & Baykal, N. (2007) MedicoPort: A medical search engine for all. *Comput. Methods Prog. Biomed.*, **86**, 73-86.
- Chatterley T., & Chojecki, D. (2010) Personal digital assistant usage among undergraduate medical students: Exploring trends, barriers and the advent of smartphones. *Journal of Medical Library Association* **98**(2), 157-160.
- Chen L., Li X., & Han J. (2013) MedRank: Discovering influential medical treatments from literature by information network analysis. In *Proceedings of the 24th Australasian Database Conference (ADC)*, Sydney, Australia (pp.3-12).

- Chowdhury, G. G. (2002) Digital libraries and reference services: present and future. *Journal of Documentation*, 58, 258-283.
- Dala-Ali BM., Lloyd MA., & Al-Abed Y. (2011) The uses of the iPhone for surgeons. *Surgeon* 9(1), 44–48. doi:10.1016/j.surge.2010.07.014.
- Dale JC., & LeFlore J. (2007). Personal digital assistants: Making the most use of them in clinical practice. *Journal of Paediatric Health Care* 21, 339-342.
- Davies K., & Harrison J. (2007) The information-seeking behaviour of doctors: A review of the evidence. *Health Information and Libraries Journal* 24(2), 78-94.
- Dee CR., Teolis M., & Todd AD. (2005) Physicians' use of the personal digital assistant (PDA) in clinical decision making. *Journal of Medical Library Association* 93(4), 480-486.
- Del Fiol G, Workman TE, Gorman PN.(2014) Clinical Questions Raised by Clinicians at the Point of CareA Systematic Review. *JAMA Intern Med.* 174(5):710-718
- Di Pietro T., Coburn G., Dharamshi N., Doran D., Mylopoulos J., Kushniruk A., Nagle L., Didani S., Tourangeau A., Laurie-Shaw B., Lefebvre N., Reid-Haughian C., Carryer J., & McArthur G. (2008). What nurses want: Diffusion of innovation. *Journal of Nursing Care Quality* 23,140-145.
- Edward Oguejiofor, Philippe Bazot ,Bruno Georges, Rebecca Huber. Callum Jackson ,Jochen Kappel, Cameron Martin, Bala S. Subramanian, Abhijit Sur. (2008) "*Developing SIP and IP Multimedia Subsystem (IMS) Applications*", 2008, Primera Edición, IBM RedBooks, Estados Unidos.146.
- Elo S., & Kyngas H. (2008) The qualitative content analysis process. *Journal of Advanced Nursing* 62(1), 107-115.
- Falagas ME., Karveli EA., & Tritsaroli VI. (2008) The risk of using the Internet as reference resource: A comparative study. *International Journal of Medical Informatics*, 77(4), 280-286.
- Franko OI., & Tirrell TF. (2012) Smartphone app use among medical providers in ACGME training programs. *J Med Syst* 36(5), 3135-9.
- Franko OI. (2011) Smartphone apps for orthopaedic surgeons. *Clin Orthop Relat Res* 469 (7), 242-248.
- Fredrik Nygard. (2015) Personolized Search. Norwegian University of Science and Technology Department of Computer Science.,phD thesis published.,Norwegeian

- Fox CR., Day D., Griffin A., & Huckstadt A. (2007) Capturing patient encounters during health care provider training using personal digital assistants. *Journal of Allied Health* **36**(4), 326-52.
- G audiat, A., Ruch, P., Joubert, M., Uziel, P., Strauss, A., Thonnet, M., Baud, R., Spahni, S., Weber, P., Bonal, J., Boyer, C., Fieschi, M. & Geissbuhler, A. (2006) Health search engine with e-document analysis for reliable search results. *International Journal of Medical Informatics*, **75**, 73-85.
- Gauch S., Speretta M., Chandramouli A., Micarelli A. (2007) User Profiles for Personalized Information Access. In: Brusilovsky P., Kobsa A., Nejdl W. (eds) *The Adaptive Web. Lecture Notes in Computer Science*, **43**(21), 54-89. Springer, Berlin, Heidelberg
- George, L. E., Davidson, L. J., Serapiglia, C. P., Barla, S. & Thotakura, A. (2010) Technology in Nursing Education: A Study of PDA Use by Students. *Journal of Professional Nursing*, **26** (14), 371-376.
- Gonzalez-Gonzalez AI., Dawes M., Sanchez-Mateos J., Riesgo-Fuertes R., Escortell-Mayor E., Sanz-Cuesta T., & Hernandez-Fernandez T. (2007) Information needs and information-seeking behaviour of primary care physicians. *Ann Fam Med* **5**(4), 345-352.
- Graham-Jones P., Jain SH., Friedman CP., Marcotte L., & Blumenthal D. (2011) The need to incorporate health information technology into physicians education and professional development. *Health Affairs* **31**(3), 481-487.
- Hanauer, D. A., Mei, Q., Law, J., Khanna, R. & Zheng, K. (2015). Supporting information retrieval from electronic health records: A report of University of Michigan's nine-year experience in developing and using the Electronic Medical Record Search Engine (EMERSE). *Journal of Biomedical Informatics*, **55**, 290-300.
- Has, M., Kaplan, A. B., Dizdaro, B. & Lu. (2015) Medical image segmentation with active contour model: Smartphone application based on client-server communication. Medical Technologies National Conference (TIPTeKNO), 2015, 15-18 Oct. 2015 2015. 1-4.
- Healthline. (2007). *Healthline*. Retrieved from <http://www.healthline.com>
- Herskovic JR., Cohen T., Subramanian D., Iyengar MS., Smith JW., & Bernstam EV. (2011) MedRank: Using graph-based concept ranking to index biomedical texts. *International Journal of Medical Informatics* **80**(6), 431-41.
- Hliaoutakis A., Zervanou K., & Petrakis E. (2009) The amtex approach in the medical document indexing and retrieval application. *Data & Knowledge Engineering* **68**(3), 380–392.

- H. Schulzrinne. The tel URI for Telephone Numbers. RFC 3966, Internet Engineering Task Force, December 2004.
- J. Rosenberg, H. Schulzrinne, G. Camarillo, A. Johnston, J. Peterson, R. Sparks, M. Handley, and E. Schooler. SIP: Session Initiation Protocol. RFC 3261, Internet Engineering Task Force, June 2002.
- Johannes L., Goeuriot L., Kelly L., & Jones GJF. (2013) DCU@TREC Med 2012: Using ad-hoc baselines for domain-specific retrieval. In *Proceedings of the Twenty-First Text REtrieval Conference (TREC 2012)*; <<http://trec.nist.gov/pubs/trec21/papers/DCU.medical.final.pdf>>
- Johnson C. (2008) Nurses and the use of personal digital assistants (PDAs) at the point of care. *Scroll* 1(1).<http://fdt.library.utoronto.ca/index.php/fdt/article/viewArticle/4906/1766>>
- Kannampallil TG., Franklin A., Mishra R., Almoosa KF., Cohen T., & Patel VL. (2013) understanding the nature of information seeking behaviour in critical care: Implications for the design of health information technology. *Artificial Intelligence in Medicine* 57(1), 21-29 Doi:10.1016/j.artmed.2012.10.002.
- Kelly, D., Teevan, J.: Implicit feedback for inferring user preference: a bibliography. *ACM SIGIR Forum* 37(2) (2003) 18-28
- Kho A., Henderson LE., Dressler DD., & Kripalani S. (2006) Use of handheld computers in medical education. a systematic review. *J Gen Intern Med.* 21(5), 531-537.
- Kohane IS., Drazen JM., & Champion EW. (2012) A glimpse of the next 100 years in medicine. *New England Journal of Medicine* 367(26), 2538-39.
- Koopman B., Bruza P., Lawley M., & Sitbon L. (2011) Evaluating medical information retrieval. In *Proceedings of SIGIR*, Beijing China (July 24-28).
- K. Rikitake, Y. Araki, Y. Kawahara, M. Minami and H. Morikawa, "NGN/IMS-based ubiquitous health monitoring system — Data format design," *2009 Sixth International Conference on Networked Sensing Systems (INSS)*, Pittsburgh, PA, 2009, pp. 1-10.
- Langton JM., Blanch B., Pesa N., Park JM., & Pearson SA. (2013) How do medical doctors use a web-based oncology protocol system? A comparison of Australian doctors at different levels of medical training using log file analysis and an online survey. *BMC Medical Informatics and Decision Making* 13(82) 20-26.
- Leino, J. (2014). 'User factors in recommender systems: Case studies in e-commerce, news recommending, and e-learning'

- Leo GD., LeRouge C., Ceriani C., & Niederman F. (2006) Websites most frequently used by physician for gathering medical information. *AMIA Annu Symp Proc* 902.
- Liang, Chunyan, "User profile for personalized web search," 2011 Eighth International Conference on Fuzzy Systems and Knowledge Discovery (FSKD), Shanghai, 2011, pp. 1847-1850
- Lindquist, A. M., Johansson, P. E., Petersson, G. I., Saveman, B.-I. & Nilsson, G. C. (2008) The Use of the Personal Digital Assistant (PDA) Among Personnel and Students in Health Care: A Review. *Journal of Medical Internet Research*, 10(4), e31. <http://doi.org/10.2196/jmir.1038>.
- Lopes CT., & Ribeiro C. (2011) Comparative evaluation of web search engines in health information retrieval. *Online Information Review* **35**(6), 869-92.
- Luo G. (2009) Design and evaluation of the iMed intelligent medical search engine. In *Proceedings of the 25th International Conference on Data Engineering*, Shanghai, March 29-April 2 (pp.1379-90).
- Luo G., Tang C., Yang H., & Wei, X. (2008) MedSearch: A specialized search engine for medical information retrieval. In *Proceedings of the 17th ACM Conference on Information and Knowledge Management*, California USA (October 26-30).
- Luo G., & Tang C. (2008) On iterative intelligent medical search. In *Proceedings of the 31st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval* (pp.3-10).
- Marcy, T. W., Kaplan, B., Connolly, S. W., Michel, G., Shiffman, R. N. & Flynn, B. S (2008) Developing a Decision Support System for Tobacco Use Counseling Using Primary Care Physicians. *Informatics in Primary Care*, **16**(2), 101–109.
- Margalit RS., Roter D., Duvenant MA., Larson S., & Reis S. (2006). Electronic medical record use and physician-patient communication: An observational study of Israeli primary care encounters. *Patient Educ Couns* **61**(3), 134-41.
- Markota, I & Cubic, Ivica. (2019). Medical Sensor Application Framework Based on IMS/SIP Platform.
- Manning, C. D., Raghavan, P. and Schütze, H. [2008], *Introduction to Information Retrieval*, Cambridge University Press, New York, NY, USA
- Mayuri Gund, Snehal Andhalkar, Prof. Dipti Patil, Dr. V.M.Wadhavi. "An Intelligent Architecture for Multi-Agent Based M-Health Care System"*International Journal of Computer Trends and Technology (IJCTT)*,V1(1):106-110 March to April Issue 2011 .ISSN 2231-2803.www.ijcttjournal.org. Published by Seventh Sense Research Group.

- McConaghy JR. (2006) Evolving medical knowledge: moving toward efficiently answering clinical questions and keeping current. *Primary Care* **33**(4), 831-837.
- McCord G., Pendleton BF., Schrop SL., Weiss L., Stockton L., & Hamrich LM. (2009) Assessing the impact on patient-physician interaction when physicians use personal digital assistants: A Northeastern Ohio Network (NEON) study. *J Am Board Fam Med* **22**(4), 353-359.
- McGowan BS., Wasko M., Vartabedian BS., Miller RS., Freiherr DD., & Abdolrasulnia M. (2012). Understanding the factors that influence the adoption and meaningful use of social media by physicians to share medical information. *J Med Internet Res* **14**(5), e117.
- McGowan JJ. (2012) Tomorrow's academic health sciences library today. *Journal of the Medical Library Association* **100**(1), 43-46.
- McKibbon K., Fridma DB., & Crowley R. (2007) How primary care physicians' attitudes towards risk and uncertainty affect their use of electronic information resources. *Journal of Medical Library Association* **95**(2), 138-146.
- Mechling, L. C., Gast, D. L. & Seid, N. H (2009) Using a Personal Digital Assistant to Increase Independent Task Completion by Students with Autism Spectrum Disorder. *Journal of Autism and Developmental Disorders*, 39, 1420-1434.
- Medline. (2010) *Medline: Worked example of a search strategy*. University of Bristol, UK. Retrieved from <www.bristol.ac.uk/library/support/subjects/medfac/searchstrategyautism.pdf>
- Medstory. (2007) *Medstory*. Retrieved from <<http://www.medstory.com>>
- MeSH.(2016). <https://www.nlm.nih.gov/mesh/meshhome.html>
- Miikka Poikselk`a and Georg Mayer, The IMS IP Multimedia Concepts and Services third edition, John Wiley & Sons Ltd 2009, ISBN 978-0-470-72196-4
- Mollon B., Chong J Jr., Holbrook Am., Sung M., Thabane L., & Foster G. (2009). Features predicting the success of computerized decision support for prescribing: A systematic review of randomized controlled trials. *BMC Med Inf Decis Mak* **9**(11).
- Montani, S. & Bellazzi, R. (2002). Supporting decisions in medical applications: the knowledge management perspective. *International Journal of Medical Informatics*, 68, 79-90.
- Moxey A., Robertson J., Newby D., Hains I., Williamson M., & Pearson SA. (2010) Computerized clinical decision support for prescribing: Provision does not guarantee uptake. *J Am Med Inform Assoc* **17**(1), 25-33.

- M. Speretta and S. Gauch, "Personalized search based on user search histories," *The 2005 IEEE/WIC/ACM International Conference on Web Intelligence (WI'05)*, Compiegne, France, 2005, pp. 622-628. doi: 10.1109/WI.2005.114
- Naidoo P., Jinabhai CC., & Taylor M. (2010) Identification of sources from which doctors in private sector obtain information on HIV and AIDS. *South African Family Practice* **52**(2), 137-141.
- Navarro, J. R. Mas, J. F. Navajas and C. P. Alcega, "Performance of a 3g-based mobile telemedicine system," *CCNC 2006. 2006 3rd IEEE Consumer Communications and Networking Conference, 2006.*, Las Vegas, NV, USA, 2006, pp. 1023-1027
- Nwagwu WE., & Oshiname R. (2009) Information needs and seeking behaviour of nurses at the University College Hospital, Ibadan, Nigeria. *African Journal of Library, Archival and Information Science* **58**, 873-881.
- Oehler RL., Smith K., & Toney JF. (2010) Infectious diseases resources for the iPhone. *Clin Infect Dis* **50**(9), 1268-1274. Doi:10.1086/651602.
- Okoro CC., & Okoro I. (2009) The use of medical information in Nigeria: The influence of gender and status. *Journal of Continuing Education in the Health Professions* **29**(4), 254-258.
- O'Malley KJ., Cook KF., Price MD., Wildes KR., Hurdle JF., & Ashton CM. (2005) Measuring diagnoses: ICD code accuracy. *Health Services Research* **40**(5pt2), 1620-39.
- OPNET. (2018). <http://opnetprojects.com/opnet-network-simulator/>
- Ozdalga E., Ozdalga A., & Ahuja N. (2012) The smartphone in medicine: A review of current and potential use among physicians and students. *Journal of Medical Internet Research* **14**(5), e128.
- Park E., & Nam HS. (2009) A service-oriented medical framework for fast and adaptive information delivery in mobile environment. *IEEE Trans Inf Technol Biomed* **13**(6), 1049-56.
- Patel VL., & Cohen T. (2008) New perspectives on error in critical care. *Current Opinion in Critical Care* **14**(4), 456-459.
- Payne, K. F. B., Wharrad, H. & Watts, K (2012) Smartphone and medical related App use among medical students and junior doctors in the United Kingdom (UK): a regional survey. *BMC Medical Informatics and Decision Making*, 12, 1-11.
- Pearson SA., Moxey A., Robertson J., Hains I., Williamson M., Reeve J., & Newby D. (2009) Do computerised clinical decision support systems for prescribing change practice? A systematic review of the literature (1990-2007). *BMC Health Serv Res* **9**(154).

- Perez BA., Von Isenburg MA., Yu M., Tuttle BD., & Adams MB. (2013) Exploring the impact of tablet computers on medical training at an academic medical center. *Journal of Medical Library Association* **101**(2), 154-157.
- Plutchak, T. S. (2012). Breaking the barriers of time and space: the dawning of the great age of librarians. *Journal of the Medical Library Association : JMLA*, 100(1), 10–19.
- Poretsky.s, J. Perser, S. Erramilli, and S. Khurana. (2009). Terminology for benchmarking network- layer traffic control mechanisms, RFC4689:11-14
- Prgomet M., Georgiou A., & Westbrook JI. (2009) The impact of mobile handheld technology on hospital physicians' work practices and patient care: a systematic review. *J Am Med Inform Assoc* **16**(6), 792–801.
- Pullen, Ravi Malghan, Lava K. Lavu, Gang Duan, Jiemei Ma, Hoon Nah. (1999) .A Simulation Model for IP Multicast with RSVP. RFC 2490: 1-31
- Ray Sunil. (2016) Essentials of Machine Learning Algorithms (with Python and R Codes). Analytics vidhya, 12May2016.
- Reddy MC., & Spence PR. (2008) Collaborative information seeking: A field study of a multidisciplinary patient care team. *Information Processing and Management* **44**(1), 242-255.
- Ricci, F., Rokach, L., Shapira, B. and Kantor, P. B., eds [2011], Recommender Systems Handbook, Springer. URL: <http://www.springerlink.com/content/978-0-387-85819-7>
- Rodriguez, N. J., Borges, J. A., Soler, Y., Murillo, V. & Sands, D. Z. (2004) A usability study of physicians' interaction with PDA and laptop applications to access an electronic patient record system. Computer-Based Medical Systems, 2004. CBMS 2004. Proceedings. 17th IEEE Symposium on, 24-25 June 2004 2004. 153-160.
- Rohini, U & Ambati, Vamshi. (2006). Improving Re-ranking of Search Results Using Collaborative Filtering. 4182. 205-216.
- RxNorm. (2016) RxNorm overview. <https://www.nlm.nih.gov/research/umls/rxnorm/overview.html>
- Salton, G. (1971). The SMART retrieval system: Experiments in automatic document processing. Prentice-Hall, Upper Saddle River, NJ,USA
- Sackett D.L. (2005) Evidence-based medicine, what it is and what it isn't. *British Medical Journal* **312**, 71-72.
- Sarwar, B., Karypis, G., Konstan, J. and Riedl, J. (2001), Item-based collaborative filtering recommendation algorithms, in 'Proceedings of the 10th international conference on World Wide Web', ACM, pp. 285–295.

- Seeger, C., Laerhoven, K. V. & Buchmann, A.(2015). MyHealthAssistant: An Event-driven Middleware for Multiple Medical Applications on a Smartphone-Mediated Body Sensor Network. *IEEE Journal of Biomedical and Health Informatics*, 19, 752-760.
- SnoMed CT. (2016). Retrieved from <https://www.nlm.nih.gov/healthit/snomedct/index.html>
- Singh V K, (2015) Vector space model: an information retrieval system. *International Journal of Advanced Engineering Research and Studies*, (12), 22-26.
- Simon Znaty and Jean-Louis Dauphin, (2005) “IP Multimedia Subsystem:Principles and Architecture”, http://www.efort.com/media_pdf/IMS_ENG.pdf
- Stephens MB., Waechter D., Williams PM., Williams AL., Yew KS., & Strayer SM. (2010) Institutional support for handheld computing: clinical and educational lessons learned. *Med Ref Serv Q* 29(1), 28–36.
- Sun Y., Yu Y., & Han J. (2009) Ranking-based clustering of heterogeneous information networks with star network schema. In *Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp.797-806).
- TREC Clinical Decision Support Track. (2014) 2014 Task: Generic clinical questions. Retrieved from <http://www.trec-cds.org/2014.html>
- Terry, M(2010) Medical Apps for Smartphones. *Telemedicine and e-Health*, 16, 17-22.
- Tumwikirize WW., Ogwal-Okeng JW., Vernby A., Anokbonggo WW., Gustafsson LL., & Lundborg CS. (2008) Access and use of medicines information sources by physicians in public hospitals in Uganda: a cross-sectional survey. *African Health Sciences* 8(4), 220-26.
- Yu, P., Decourten, M., Pan, E., Galea, G. & Pryor, J. (2009). The development and evaluation of a PDA-based method for public health surveillance data collection in developing countries. *International Journal of Medical Informatics*, 78, 532-542.
- Vlveness LV., Aasland O., Nylenna M. & Kristiansen I.S. (2009) Norwegian physicians’ knowledge of and opinions about evidence-based medicine: Cross-sectional study. *Plos One* 4(1), e7828.
- Vollmar HC., Rieger MA., Butzlaff ME., & Ostermann T. (2009) General practitioners preferences and use of educational media: A German perspective. *BMC Health Services Research* 9(31). Doi:10.1186/1472-6963-9-31

- Wang, L., Wang, J., Wang, M., Li, Y., Liang, Y., Xu, D (2012). Using Internet search engines to obtain medical information: a comparative study. *us national library of medicine*, 14, 74.
- Westbrook JL., Coiera EW., & Gosling AS. (2005) Do online information retrieval systems help experienced clinicians answer clinical questions? *Journal of the American Medical Informatics Association* **12**(3): 315-321.
- WHO. (2015). ICD <http://www.who.int/classifications/icd/en/>
- Wu, I.-L., Li, J.-Y. & Fu, C.-Y. (2011) The adoption of mobile healthcare by hospital's professionals: An integrative perspective. *Decision Support Systems*, 51, 587-596.
- Zhang, Y., Wang, P., Heaton, A. & Winkler, H. (2012). Health information searching behaviour in MedlinePlus and the impact of tasks. *Proceedings of the 2nd ACM SIGHIT International Health Informatics Symposium*. Miami, Florida, USA: ACM.